

T A / Service
N70-25226
N70-25246

NATIONAL AERONAUTICS AND SPACE ADMINISTRATION

NASA CR-109503

Space Programs Summary 37-60, Vol. III

Supporting Research and Advanced Development

For the Period October 1 to November 30, 1969

CASE FILE
COPY

JET PROPULSION LABORATORY
CALIFORNIA INSTITUTE OF TECHNOLOGY
PASADENA, CALIFORNIA

December 31, 1969

NATIONAL AERONAUTICS AND SPACE ADMINISTRATION

Space Programs Summary 37-60, Vol. III

Supporting Research and Advanced Development

For the Period October 1 to November 30, 1969

JET PROPULSION LABORATORY
CALIFORNIA INSTITUTE OF TECHNOLOGY
PASADENA, CALIFORNIA

December 31, 1969

SPACE PROGRAMS SUMMARY 37-60, VOL. III

Copyright © 1970
Jet Propulsion Laboratory
California Institute of Technology

Prepared Under Contract No. NAS 7-100
National Aeronautics and Space Administration

Preface

The Space Programs Summary is a multivolume, bimonthly publication that presents a review of technical information resulting from current engineering and scientific work performed, or managed, by the Jet Propulsion Laboratory for the National Aeronautics and Space Administration. The Space Programs Summary is currently composed of four volumes.

- Vol. I. *Flight Projects* (Unclassified)
- Vol. II. *The Deep Space Network* (Unclassified)
- Vol. III. *Supporting Research and Advanced Development* (Unclassified)
- Vol. IV. *Flight Projects and Supporting Research and Advanced Development* (Confidential)

Contents

ADVANCED STUDIES DIVISION

I. Future Projects	1
A. Imaging and Sampling Requirements for an Automated Lunar Roving Vehicle	
<i>R. G. Brereton, G. E. Ulrich, and D. H. Dahlem, NASA Code 195-06-10-01</i>	1

DATA SYSTEMS DIVISION

II. Computation and Analysis	4
A. The Numerical Solution of Two-Point Boundary Value Problems With Linear Boundary Constraints	
<i>J. P. Keener, NASA Code 129-04-20-03</i>	4
B. On the Design of a Primitive Fast Fourier Transformation Process	
<i>F. T. Krogh, NASA Code 129-04-20-03</i>	9

SPACE SCIENCES DIVISION

III. Space Instruments	14
A. A Spherical Coil System for the Helium Vector Magnetometer	
<i>W. L. Momsen, NASA Code 188-36-01-04</i>	14
IV. Lunar and Planetary Sciences	18
A. Effect of Fe, Ca, and Mn Concentrations on Thermoluminescence Intensity of Meteoritic Enstatites	
<i>R. T. Greer, NASA Code 195-42-12-01</i>	18
V. Physics	20
A. The de Haas—van Alphen Effect in Indium	
<i>R. W. Vaughan and D. D. Elleman, NASA Code 129-02-21-04</i>	20
B. Integral Theorems in Dyadic Notation	
<i>H. D. Wahlquist, NASA Code 129-04-20-04</i>	22
C. Collisional Quenching of $O(^1D)$	
<i>W. B. DeMore, NASA Code 129-01-20-02</i>	26

TELECOMMUNICATIONS DIVISION

VI. Communications Elements Research	28
A. Spacecraft Antenna Research: Fields in a Cylindrically Inhomogeneous Plasma Excited by an Arbitrarily Oriented Dipole Antenna	
<i>R. Woo, NASA Code 125-21-09-07</i>	28

Contents (contd)

B. Spacecraft Antenna Research: High-Efficiency S- and X-Band Telemetry and Tracking Feed	
K. Woo, NASA Code 186-68-04-27	36
VII. Spacecraft Telemetry and Command	42
A. Telemetry Data Systems Implementation	
L. A. Couvillon, NASA Code 186-68-04-29	42
VIII. Spacecraft Telecommunications Systems	46
A. Data-Aided Carrier Tracking Loops	
W. C. Lindsey and M. K. Simon, NASA Code 125-21-09-06	46
B. A Comparison of the Performance of Costas or Squaring Loops With Data-Aided Loops	
W. C. Lindsey and M. K. Simon, NASA Code 125-21-09-06	56
C. Calculations of Power Spectra of Square-Wave Modulated Pseudorandom Sequences	
J. H. Yuen, NASA Code 160-21-12-03	58

GUIDANCE AND CONTROL DIVISION

IX. Spacecraft Power	62
A. Photovoltaic Radiation Effects	
P. A. Berman, NASA Code 120-33-14-04	62
B. Photovoltaic Supporting Development	
R. K. Yasui, NASA Code 120-33-13-03	63
C. A Circuit Design for Eliminating High Current Spikes at Converter and Inverter Inputs	
C. McLyman, NASA Code 120-60-04-01	66
D. Quad Redundant Shunt Regulator for Nuclear Power Subsystem	
D. Hopper and H. Wick, NASA Code 120-60-04-03	69
E. Battery Storage Optimization and Design Studies	
R. S. Bogner and R. E. Patterson, NASA Code 120-34-10-15	73
F. Impedance of Silver-Zinc Battery Cells	
W. L. Long, NASA Code 120-34-10-15	76
G. Heat Sterilizable Battery Development	
R. Lutwack, NASA Code 120-34-10-03	79
H. Performance Testing of the SNAP-27 Thermoelectric Generator	
G. Stapfer, NASA Code 120-27-41-04	80
I. The Application of Bench Tests in the Development of Heat Sterilizable Battery Separators	
W. von Hartmann, NASA Code 120-34-10-05	84

Contents (contd)

J. Long-Term Life Test of the Spare Mariner Venus 67 Power Subsystem Hardware	
A. Krug, NASA Code 120-60-04-02	88
K. Interactions Between Radiation Fields From RTGs and Scientific Experiments on Spacecraft	
C. G. Miller and V. C. Truscello, NASA Code 120-27-41-06	88
X. Guidance and Control Analysis and Integration	99
A. Orbit Insertion Maneuver Mechanization Study	
G. Jaivin, NASA Code 125-17-15-05	99
XI. Spacecraft Control	103
A. Digital Sun Sensor	
L. F. Schmidt, NASA Code 186-68-02-45	103
B. Stability and Control of the Solar-Electric Ion-Engine Propelled Spacecraft	
L. L. Schumacher, NASA Code 120-26-16-07	107
C. A Simple Digital-Derived-Rate Scheme	
H. S. Lin, NASA Code 186-68-02-41	110
D. Nonlinear Control System Design Analysis via Nonlinear Filtering Techniques	
A. K. Bejczy, NASA Code 125-17-15-10	112
E. Inertial Sensor Sterilization	
P. J. Hand, NASA Code 186-68-02-44	119
XII. Guidance and Control Research	121
A. Cesium Breakdown Voltage in Thermionic Diodes	
K. Shimada, NASA Code 129-02-22-07	121
B. Narrow Curie-Point Switching Transfer Characteristics in MnBi Films	
G. Lewicki and J. Guisinger, NASA Code 129-02-21-06	124

ENGINEERING MECHANICS DIVISION

XIII. Materials	127
A. Tape-Type Temperature Control Coatings Application	
J. C. Lewis, NASA Code 124-09-26-05	127
B. Sterilization and Thermal-Vacuum Effects on Spacecraft Polymeric Materials	
W. Roper, NASA Code 186-68-13-09	136
XIV. Applied Mechanics	142
A. High-Impact Dynamic-Response Analysis of Nonlinear Structures	
K. K. Gupta, NASA Code 120-34-10-03	142

Contents (contd)

B. Eigenvalue of $(\mathbf{B} - \lambda \mathbf{A}^*)\mathbf{y} = 0$ With Positive Definite Band Symmetric \mathbf{B} and Band Hermitian \mathbf{A}^* and Its Application to Natural Frequency Analysis of Flexible Space Vehicles <i>K. K. Gupta, NASA Code 124-08-26-02</i>	145
C. Nonstationary Peak Distribution <i>M. Shinozuka and J. N. Yang, NASA Code 124-08-26-02</i>	149
D. An Analytical and Experimental Study of Heat Transfer in a Simulated Martian Atmosphere <i>D. Ting, NASA Code 124-09-18-06</i>	153
E. Development of Advanced Composite Liquid Propulsion Pressure Vessels <i>A. Knoell, NASA Code 731-12-42-01</i>	156
F. Optimization of Space Antenna Concepts <i>E. Heer and J. N. Yang, NASA Code 124-08-26-01</i>	157
G. A Frequency Domain Solution for the Linear Attitude Control Problem of Spacecraft With Flexible Appendages <i>M. R. Trubert and J. A. Garba, NASA Code 120-33-16-04</i>	159
H. Cabling Heat Loss at Neptune <i>L. D. Stimpson, NASA Code 124-09-26-06</i>	163
I. Gamma Ray Spectrometer Preliminary Thermal Tests <i>S. W. Petrick, NASA Code 914-49-10-02</i>	166
XV. Electronics Parts Engineering	174
A. Matrix Test of Sterilizable Piece Parts <i>K. Martin, NASA Code 186-58-13-08</i>	174
XVI. Advanced Projects Development	177
A. TOPS Boom Deployment Mechanisms <i>E. E. Sabelman, NASA Code 186-68-12-05</i>	177

PROPULSION DIVISION

XVII. Solid Propellant Engineering	181
A. A Statistical Model for Radiation Penetration in Composite Propellants <i>R. L. Klaus, NASA Code 128-32-90-01</i>	181
B. Heat Conduction in a Semi-Infinite Solid With a Moving Surface With Application to Solid Propellant Rocket Combustion <i>R. L. Klaus, NASA Code 128-32-90-01</i>	187
C. Functionality Determination of Binder Prepolymers by Thin Layer Chromatography <i>H. E. Marsh, Jr., and W. M. Foley, Jr., NASA Code 128-32-80-02</i>	199

Contents (contd)

D. Gelation Studies	
<i>H. E. Marsh, Jr., and J. F. Wolfe, NASA Code 128-32-80-02</i>	200
E. Electronic Conductivity of Solid Propellant Exhaust Products	
<i>O. K. Heiney, NASA Code 128-32-90-03</i>	202
XVIII. Research and Advanced Concepts	206
A. Closed-Loop Operation of a Hollow-Cathode Ion Thruster	
<i>E. V. Pawlik and T. D. Masek, NASA Code 120-26-16-01</i>	206
XIX. Liquid Propulsion	211
A. Material Compatibility	
<i>L. R. Toth and O. F. Keller, NASA Code 731-13-44-02</i>	211
B. Advanced Injector Fabrication Techniques	
<i>R. W. Riebling, NASA Code 731-12-42-04</i>	212

MISSION ANALYSIS DIVISION

XX. Systems Analysis Research	216
A. Relativistic Solar Perturbations on the Moon in the Brans-Dicke Theory	
<i>J. D. Anderson, NASA Code 129-04-22-03</i>	216
B. Propagation of Statistical Moments in Nonlinear Systems and the Application to a Planetary Orbiter	
<i>J. F. Jordan, NASA Code 120-26-16-03</i>	217
C. The Statistics of the Solution of a Nonlinear Differential Equation With Random Initial Conditions	
<i>H. Lass, NASA Code 129-04-20-02</i>	220
D. An n Thruster Reliability Problem for Electric Propulsion	
<i>W. Kizner, NASA Code 120-26-16-03</i>	223

I. Future Projects

ADVANCED STUDIES DIVISION

A. Imaging and Sampling Requirements for an Automated Lunar Roving Vehicle, R. G. Brereton, G. E. Ulrich, and D. H. Dahlem

1. Introduction

A field program to determine the practicability of, and examine some science requirements for, performing geologic tasks from an automated roving vehicle was described in SPS 37-55, Vol. III, pp. 258-263. This article represents a continued review of the field data, as supplemented by discussions with interested scientists and engineers and members of the Joint Panel for Lunar Mobile Science and Traverse Planning.¹ It is specifically oriented toward defining the science requirements for the imaging, surface sampling, and sample storage systems for an automated lunar roving vehicle.

2. Science-Related Imaging Requirements

The imaging capability must encompass these primary functions:

- (1) Geologic reconnaissance.
- (2) Site selection for rock and soil samples and for instrument deployment.

- (3) Observation and direction of sampling operation at a visual resolution of shapes of particles at 0.5 mm at point of collection.
- (4) Examination of sample.
- (5) Terrain assessment for navigation, guidance, and topographic analysis.

For a user remotely located from the scene, the images received must be essentially as clear as those provided by the eyes of a geologist or careful observer on the site. This requires a camera system that can record a panoramic view such that the near and the far field are integrated into a full field with normal perspective. This camera system should be capable of resolving the textural detail of rock or soil fragments in the near field (~ 1 m) surface that are as small as 0.5 mm in diameter.

In general, scientific observations from an automated rover on the lunar surface require:

- (1) The observing capability equal to that of a human observer using his unaided eye for panoramic viewing, using a hand-held 10-power magnifying glass for specimen analysis, and finally using a 10-power telescope for viewing select portions of the panorama.

¹Sponsored by George C. Marshall Space Flight Center.

- (2) Panoramic (reconnaissance) viewing with widest possible latitude in horizontal field of view and about 70 deg in vertical field of view.
- (3) Capability to vary resolution, depth of field, field of view, and frame time.
- (4) Capability to vary lighting conditions by rotating objects in near field along two perpendicular axes, changing camera orientation, and artificial illumination.
- (5) Capability to reduce selected frames photogrammetrically, which requires stereo imagery (0.1 to 1.0-m based image pairs), metric quality imagery or imagery which can be reconstituted on earth to metric quality, and position and orientation of camera with respect to lunar coordinates and with respect to local vertical (or horizontal).
- (6) Camera orientation with respect to vehicle heading and local vertical. Color vision could be important locally for rock and mineral identification. Reconstituted color may be satisfactory for image enhancement and field examinations.

Some technique of real-time ranging to observed features must be worked out. Stereo viewing may be practical in the near field, but radar or laser ranging may be required to relate distances over the panoramic view to ± 10 m.

3. Sampling Requirements

The sampler must be able to fulfill the following requirements:

- (1) Coordination of imaging with sampling activities.
- (2) Sampling of particulate material.
- (3) Sampling of large rock fragments.
- (4) Preliminary assessment of sampled material for texture, mineralogy, and chemistry.
- (5) Storage of selected samples.

a. Sample collection. A sampling instrument similar in function and design to the *Surveyor* sampler and in contrast to a drill or auger-type sampler appears to have the most utility for collection of particulate material. This type of sampler can be directed to any select fragment or area on the surface, scratch at the surface of outcrops, uncover contacts, trench to permit a TV view of the subsurface, and provide data about the vertical distribution of surface material.

It is also necessary to have a device capable of coring, chipping, or breaking fragments from large rocks or outcrops. This type of whole sample for earth return is preferable to a comminuted one in which the textural properties are largely destroyed. A core of approximately 1 by 2 in. (26 cm³) is required, and the orientation of the sample should be known.

For the long traverse reconnaissance-type mission in which no samples are returned, a drilled sample of cuttings from coherent surface rock is preferable to a chipped or cored one. The acquired samples must be analyzed by a variety of fixed instruments, and a drill tip can provide the size of fragments required for analysis without additional sample grinding or sorting. A drill-type sampler also provides a low-power rapid technique for subsurface sampling, and it should be designed for use against vertical or sloping outcrops.

An auger-type sampler may have some practicability on a rendezvous mission for digging up a pile of subsurface material that can be subsequently picked over, sorted, and selected by the *Surveyor*-type sampler. This type of operation must be viewed by the near-field imaging system. Analysis of the chemical and mineralogical composition of sampled material prior to storage is preferable.

b. Sample preparation. To obtain reliable and accurate data from several of the analytical instruments currently proposed, it is desirable to prepare the sample before analysis. The intent in sample preparation is to allow the instrument to generate the best data possible and not to constrain the accuracy or precision of the data by the condition of the sample. It is important to include sample preparation as a factor in the trade-off studies involving expected science return, power, weight, and volume. The preparation, in general, will involve one or more of the operations of comminuting, shaping, and containing the sample. The following is not an exhaustive list of sample conditioning activities, but is representative of the types of procedures necessary:

X-ray diffraction. The sample should be particulate and fine-grained ($< 50 \mu\text{m}$). It must be presented to the instrument in a fixed form either on an adhesive backing or in a sample tray. The volume required is on the order of 1 cm³ or less.

X-ray spectrometry. The sample should be particulate, fine-grained, and either less than 20 μm in grain size or less than 50 to 60 μm with no more than a 20% range

in particle size. The sample surface must be presented in a fixed form, either in a container or restrained in shape by an instrument-sample interface.

Neutron activation. The sample volume subject to activation must be constant or known.

Microscopic viewing in polarized light. The sample (either particulate or coherent) must have portions less than 50 μm in thickness. It must be embedded in or held by a mount in such a way as to allow transmitted light viewing and positioning in the imaging system.

c. Sample storage. For the lunar sample return mission in which the automated rover makes a rendezvous with a manned mission, separate sample storage will be required for particulate material and for whole rocks.

The particulate storage device will have to, in addition to storing samples, perform the following functions:

- (1) Store a specified number of samples of lunar surface particulate material. The total number of samples will be dictated by the amount of weight allocated to this part of the mission.
- (2) Hold core samples with cuttings as well as particulate samples. This will require a volume of at least 50 cm^3 .

- (3) Present collected samples to an array of chemical, mineralogical and texture analysis instruments. Keep or reject any sample after analysis, send any sample directly to storage without analysis, or recall any sample.
- (4) Identify samples by storage position and by visual examination of indexed numbers on sample containers.
- (5) Package stored samples to minimize intercontamination and loss or transfer of material and to facilitate extraction of sample-containing modules for return to earth (hermetic sealing is not a requirement during traverse).
- (6) Retrieve and dump stored samples, probably included with containers, to prevent contamination of higher priority samples found later.

In addition to a particulate sample handling and storage device, a container should be provided that is accessible from the sample acquisition device for depositing rock fragments that will not fit into the particulate storage mechanism. It is desirable to bag, or at least separate, these fragments to avoid any intercontamination, and also to index them for subsequent identification. The camera system could be used to identify each rock sample on the basis of factors such as size, shape, number of facets, and color.

II. Computation and Analysis

DATA SYSTEMS DIVISION

A. The Numerical Solution of Two-Point Boundary Value Problems With Linear Boundary Constraints, *J. P. Keener*

1. Boundary Value Problems

The general two-point boundary value problem with linear constraints can be posed as follows:

Find $\mathbf{u}(t)$, satisfying

$$\mathbf{z} = \mathbf{f}(t, \mathbf{u}), \quad t \in [a, b] \quad (1)$$

which renders

$$\|A\mathbf{u}(a) + B\mathbf{u}(b) - \gamma\| \quad (2)$$

a minimum, where

$$\mathbf{z} = \begin{bmatrix} y_1^{(\mu_1)} \\ y_2^{(\mu_2)} \\ \cdot \\ \cdot \\ \cdot \\ y_v^{(\mu_v)} \end{bmatrix}, \quad \mathbf{f} = \begin{bmatrix} f_1(t, \mathbf{u}) \\ f_2(t, \mathbf{u}) \\ \cdot \\ \cdot \\ \cdot \\ f_v(t, \mathbf{u}) \end{bmatrix}, \quad \mathbf{u} = \begin{bmatrix} y_1 \\ y_1' \\ \cdot \\ \cdot \\ \cdot \\ y_1^{(\mu_1-1)} \\ y_2 \\ y_2' \\ \cdot \\ \cdot \\ \cdot \\ y_2^{(\mu_2-1)} \\ y_3 \\ \cdot \\ \cdot \\ \cdot \\ y_v^{(\mu_v-1)} \end{bmatrix}$$

$y_i^{(\mu_i)}$ denotes the μ_i th derivative of y_i with respect to t , A and B are $m \times n$ matrices with

$$n = \sum_{i=1}^v \mu_i, \quad m \geq n$$

γ is an m -vector, and $\|\cdot\|$ denotes the Euclidean (ℓ_2) norm of a vector. Note that if $m = n$, Expression (2) can be written as

$$A\mathbf{u}(a) + B\mathbf{u}(b) = \gamma$$

Thus, the problem involves a system of v differential equations, the i th equation of which is of order μ_i , which is solved in such a way as to satisfy the m -linear boundary conditions in a "least-squares" sense.

There are two general approaches to finding a numerical solution of this problem. One can take values which minimize Expression (2) and an "approximate" solution of Eq. (1) and then iterate in such a way as to improve the solution of Eq. (1). Alternately, one can start with a solution of Eq. (1) and an approximate solution of Expression (2), and then iterate so as to minimize Expression (2). The procedure used here is more closely related to the latter method, although aspects of both approaches have been incorporated.

The second of these approaches, known as the "shooting method," assumes that corresponding to the solution of the given problem, there exists an n -vector \mathbf{s} such that the solution $\mathbf{u}(\mathbf{s};t)$ of the initial value problem

$$\mathbf{z} = \mathbf{f}(t, \mathbf{u}), \quad \mathbf{u}(a) = \mathbf{s}, \quad t \in [a, b] \quad (3)$$

minimizes

$$\|\Phi(\mathbf{s})\| = \|A\mathbf{s} + B\mathbf{u}(\mathbf{s};b) - \gamma\| \quad (4)$$

Thus, given an approximate solution \mathbf{s}^0 , there exists a $\Delta\mathbf{s}^0$ such that

$$\|\Phi(\mathbf{s}^0 + \Delta\mathbf{s}^0)\|$$

is minimized. Expanding in a Taylor series about \mathbf{s}^0 gives that

$$\|\Phi(\mathbf{s}^0) + \frac{\partial\Phi}{\partial\mathbf{s}} \Delta\mathbf{s}^0 + \theta((\Delta\mathbf{s}^0)^2)\|$$

is minimized.

By dropping higher order terms, a natural iteration scheme suggests itself. Let \mathbf{s}^0 be an initial guess. Define a sequence $\{\mathbf{s}^v\}$ by

$$\mathbf{s}^{v+1} = \mathbf{s}^v + \Delta\mathbf{s}^v \quad (5)$$

where $\Delta\mathbf{s}^v$ minimizes

$$\left\| \Phi(\mathbf{s}^v) + \frac{\partial\Phi}{\partial\mathbf{s}} \Delta\mathbf{s}^v \right\| \quad (6)$$

This approach, although quite good for some boundary value problems, is totally inadequate for a general subroutine. Suppose, for instance, that a dominant solution of Eq. (1) completely overwhelms the solution which minimizes Expression (2). Or, suppose that for initial conditions \mathbf{s}^0 , there is a singularity in the interval $[a, b]$, or that a solution to the initial value problem does not exist. In such cases, it is impossible to evaluate $\Phi(\mathbf{s}^v)$ with any significance whatsoever.

To overcome this difficulty, we introduce a variant of shooting known as multiple or parallel shooting (Refs. 1 and 2), which is designed to reduce the growth of errors in the solutions of the given initial value problems. This is done by introducing points $a = t_0 < t_1 < \dots < t_{J-1} < t_J = b$, integrating appropriate initial value problems over each subinterval $[t_i, t_{i+1}]$, and then simultaneously adjusting the initial data to satisfy the boundary conditions and appropriate continuity conditions at the breakpoints.

The problem can now be formulated in the following way: Find $\mathbf{u}_i(\xi_i; t)$ satisfying

$$\mathbf{z}_i = \mathbf{f}(t, \mathbf{u}_i), \quad \mathbf{u}_i(t_{i-1}) = \xi_i, \quad t \in [t_{i-1}, t_i], \quad i = 1, 2, \dots, J \quad (7)$$

which renders

$$\|\Phi(\mathbf{u}_1(t_0), \mathbf{u}_2(t_1), \dots, \mathbf{u}_J(t_{J-1}))\| \quad (8)$$

a minimum, where

$$\begin{aligned} \Phi &= \begin{bmatrix} \mathbf{r}_B \\ \mathbf{r}_C \end{bmatrix} \\ \mathbf{r}_B &= A\mathbf{u}_1(a) + B\mathbf{u}_J(b) - \gamma \\ \mathbf{r}_C &= \begin{bmatrix} \mathbf{u}_2(t_1) - \mathbf{u}_1(t_1) \\ \mathbf{u}_3(t_2) - \mathbf{u}_2(t_2) \\ \vdots \\ \mathbf{u}_J(t_{J-1}) - \mathbf{u}_{J-1}(t_{J-1}) \end{bmatrix} \end{aligned}$$

Apparently, \mathbf{r}_B is the residual associated with the boundary conditions, and \mathbf{r}_C is the residual associated with the continuity conditions at the points t_1, t_2, \dots, t_{J-1} .

Let ξ be the nJ -vector

$$\xi = \begin{bmatrix} \mathbf{u}_1(t_0) \\ \mathbf{u}_2(t_1) \\ \vdots \\ \mathbf{u}_J(t_{J-1}) \end{bmatrix}$$

and let ξ^0 be an initial guess, then in the same manner as before, define a sequence $\{\xi^v\}$ where

$$\xi^{v+1} = \xi^v + \Delta \xi^v$$

such that $\Delta \xi^v$ minimizes

$$\|\Phi(\xi^v) + \left(\frac{\partial \Phi}{\partial \xi} \Big|_{\xi=\xi^v} \right) \Delta \xi^v\| \quad (9)$$

To understand the structure of the Jacobian matrix

$$\frac{\partial \Phi}{\partial \xi} \Big|_{\xi=\xi^v}$$

in Expression (9) or $\partial \Phi / \partial \mathbf{s}$ in Expression (6), let

$$\xi_i = \mathbf{u}_i(t_{i-1}), \quad i = 1, 2, \dots, J$$

so that

$$\xi = \begin{bmatrix} \xi_1 \\ \xi_2 \\ \vdots \\ \xi_J \end{bmatrix}$$

and let

$$\mathbf{d}_i = \mathbf{u}_{i+1}(t_i) - \mathbf{u}_i(t_i), \quad i = 1, 2, \dots, J-1$$

so that

$$\mathbf{r}_C = \begin{bmatrix} \mathbf{d}_1 \\ \mathbf{d}_2 \\ \vdots \\ \mathbf{d}_{J-1} \end{bmatrix}$$

Let

$$W_i(t) = \frac{\partial \mathbf{u}_i}{\partial \xi_i}(t)$$

Then $W_i(t)$ satisfies the matrix equation

$$\frac{\partial \mathbf{z}_i}{\partial \xi_i} = \frac{\partial \mathbf{f}}{\partial \mathbf{u}} W_i(t), \quad W_i(t_{i-1}) = I, \quad t \in [t_{i-1}, t_i] \quad (10)$$

where

$$\frac{\partial \mathbf{f}}{\partial \mathbf{u}} = \begin{bmatrix} \frac{\partial f_1}{\partial y_1} & \cdot & \cdot & \cdot & \frac{\partial f_1}{\partial y_v^{(\mu_v-1)}} \\ \frac{\partial f_2}{\partial y_1} & \cdot & \cdot & \cdot & \cdot \\ \cdot & & & & \cdot \\ \cdot & & & & \cdot \\ \frac{\partial f_v}{\partial y_1} & \cdot & \cdot & \cdot & \frac{\partial f_v}{\partial y_v^{(\mu_v-1)}} \end{bmatrix}$$

$$\frac{\partial \mathbf{z}}{\partial \xi_i} = \begin{bmatrix} \left(\frac{\partial y_1}{\partial \xi_{i,1}} \right)^{(\mu_1)} & \left(\frac{\partial y_1}{\partial \xi_{i,2}} \right)^{(\mu_1)} & \cdot & \cdot & \cdot & \left(\frac{\partial y_1}{\partial \xi_{i,n}} \right)^{(\mu_1)} \\ \left(\frac{\partial y_2}{\partial \xi_{i,1}} \right)^{(\mu_2)} & \cdot & \cdot & \cdot & \cdot & \left(\frac{\partial y_2}{\partial \xi_{i,n}} \right)^{(\mu_2)} \\ \vdots & & & & & \vdots \\ \left(\frac{\partial y_v}{\partial \xi_{i,1}} \right)^{(\mu_v)} & \cdot & \cdot & \cdot & \cdot & \left(\frac{\partial y_v}{\partial \xi_{i,n}} \right)^{(\mu_v)} \end{bmatrix}$$

and $\partial \mathbf{z}_i / \partial \xi_i$ is $\partial \mathbf{z} / \partial \xi_i$ on the i th subinterval $[t_{i-1}, t_i]$.

$W_i(t)$ is thus the variational matrix associated with \mathbf{u}_i . Letting W_i denote $W_i(t_i)$, it is clear that

$$\frac{\partial \mathbf{r}_B}{\partial \xi_i} = \begin{cases} A, & i = 1 \\ 0, & 1 < i < J \\ BW_J, & i = J \end{cases}$$

and

$$\frac{\partial \mathbf{d}_i}{\partial \xi_j} = \begin{cases} -W_i, & i = j \\ I, & i = j - 1 \\ 0, & \text{otherwise} \end{cases}$$

Piecing this together gives

$$\frac{\partial \Phi}{\partial \xi} = \begin{bmatrix} A & 0 & \cdot & \cdot & \cdot & BW_J \\ -W_1 & I & \cdot & \cdot & \cdot & 0 \\ 0 & -W_2 & I & \cdot & \cdot & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & \cdot & \cdot & \cdot & -W_{J-1} & I \end{bmatrix} \quad \text{for } J > 1 \quad (11)$$

and

$$\frac{\partial \Phi}{\partial \xi} = \frac{\partial \Phi}{\partial \mathbf{s}} = (A + BW_1) \quad \text{for } J = 1 \quad (12)$$

2. Eigenvalue Problems

The introduction of eigenvalues into multiple shooting formulation is a simple matter. In the notation of Expressions (7) and (8), seek a solution $\mathbf{u}_i(\xi_i, \lambda, t)$ satisfying

$$\begin{aligned} \mathbf{z}_i &= \mathbf{f}(t, \mathbf{u}_i, \lambda), & \mathbf{u}_i(t_i) &= \xi_i, & t &\in [t_{i-1}, t_i], \\ & & i &= 1, 2, \dots, J \end{aligned} \quad (13)$$

and minimizing

$$\begin{aligned} &\|\Phi(\mathbf{u}_1(t_0), \mathbf{u}_2(t_1), \dots, \mathbf{u}_J(t_{J-1}), \lambda)\| \equiv \\ &\|\Phi(\xi_1, \xi_2, \dots, \xi_J, \lambda)\| \end{aligned} \quad (14)$$

where λ is a q -vector of eigenvalues

It is important to realize that, without imposing proper normalization or orthogonality conditions, problems of this sort are often improperly posed. For example, in the problem $y'' = -\lambda^2 y$ (which will be discussed in detail later) possible normalization conditions are

$$\int_0^1 y_i^2 d\tau = 1$$

where y_i is the i th eigenfunction, and orthogonality conditions appear as

$$\int_0^1 y_i y_j d\tau = 0, \quad i \neq j$$

These conditions can be set up as additional differential equations with appropriate conditions. For example, letting

$$v'(t) = y_i^2(t), \quad v(0) = 0$$

implies that

$$v(t) = \int_0^t y_i^2(\tau) d\tau$$

which finally gives the condition $v(1) = 1$. Similarly, letting

$$w_{ij}(t) = \int_0^t y_i(\tau) y_j(\tau) d\tau, \quad i \neq j$$

gives that

$$w'_{ij} = y_i y_j, \quad w(0) = 0, \quad w(1) = 0$$

By incorporating conditions of this type into Expressions (13) and (14), the problem becomes properly posed.

Letting ξ be the nJ -vector of initial values, let ξ^0 and λ^0 be given initial guesses to the solution of Expressions (13) and (14). Define a sequence $\{\xi^v, \lambda^v\}$ by

$$\xi^{v+1} = \xi^v + \Delta \xi^v, \quad \lambda^{v+1} = \lambda^v + \Delta \lambda^v \quad (15)$$

where $\Delta \xi^v$ and $\Delta \lambda^v$ minimize

$$\left\| \Phi(\xi^v, \lambda^v) + \left(\frac{\partial \Phi}{\partial \xi} \right)_{\xi=\xi^v} \Delta \xi^v + \left(\frac{\partial \Phi}{\partial \lambda} \right)_{\lambda=\lambda^v} \Delta \lambda^v \right\| \quad (16)$$

or

$$\left\| \Phi(\xi, \lambda^v) + \left[\frac{\partial \Phi}{\partial \xi} \right]_{\xi=\xi^v}, \quad \frac{\partial \Phi}{\partial \lambda} \right|_{\lambda=\lambda^v} \left[\begin{matrix} \Delta \xi^v \\ \Delta \lambda^v \end{matrix} \right] \right\| \quad (17)$$

The form

$$\frac{\partial \Phi}{\partial \xi} \Big|_{\xi=\xi^v}$$

is given by Eq. (11). Let

$$V_i = \frac{\partial \mathbf{u}_i}{\partial \lambda}(t_i)$$

be the $n \times q$ variational matrix associated with λ at t_i .

Then, with

$$\Phi = \begin{bmatrix} \mathbf{r}_B \\ \mathbf{r}_C \end{bmatrix}$$

it follows that

$$\mathbf{r}_B = A\mathbf{u}_1(a) + B\mathbf{u}_J(b) - \gamma$$

$$\mathbf{r}_C = \begin{bmatrix} \mathbf{u}_2(t_1) - \mathbf{u}_1(t_1) \\ \vdots \\ \mathbf{u}_J(t_{J-1}) - \mathbf{u}_{J-1}(t_{J-1}) \end{bmatrix} = \begin{bmatrix} \xi_1 - \mathbf{u}_1(t_1) \\ \vdots \\ \xi_J - \mathbf{u}_{J-1}(t_{J-1}) \end{bmatrix}$$

$$\frac{\partial \mathbf{r}_B}{\partial \lambda} = BV_J$$

and

$$\frac{\partial \mathbf{r}_C}{\partial \lambda} = \begin{pmatrix} -V_1 \\ \vdots \\ -V_{J-1} \end{pmatrix}$$

so that, quite simply,

$$\left. \frac{\partial \Phi}{\partial \lambda} \right|_{\lambda=\lambda^0} = \begin{pmatrix} BV_J \\ -V_1 \\ \vdots \\ -V_{J-1} \end{pmatrix}$$

3. Parameter Studies

Frequently in applications, the presence of certain parameters within a boundary value problem makes it desirable to obtain solutions for a range of parameter values. The method which is used here, called Poincare continuation, consists of finding initial guesses ξ_j^0, λ_j^0 for the j th problem, from the solutions ξ_{j-1}, λ_{j-1} of the $j-1$ st problem. In the notation of Expressions (13) and (14), seek a solution $u_i(\xi_i, \lambda, \sigma; t)$ satisfying

$$\mathbf{z}_i = \mathbf{f}(t, \mathbf{u}_i, \lambda, \sigma), \quad \mathbf{u}_i(t_i) = \xi_i, \quad t \in [t_{i-1}, t_i],$$

$$i = 1, 2, \dots, J$$

(18)

and minimizing

$$\|\Phi(\mathbf{u}_1(t_0), \mathbf{u}_2(t_1), \dots, \mathbf{u}_J(t_{J-1}), \lambda, \sigma)\| = \|\Phi(\xi, \lambda, \sigma)\|$$

(19)

where σ is a p -vector of parameters. Eigenvalues need not be present in Eqs. (18) and (19), although the analysis is essentially the same in both cases.

Let the values of σ for K problems be given by $\sigma_j, j = 1, 2, \dots, K$, and let $\sigma_{j+1} = \sigma_j + \Delta \sigma_j$. Then, for the $j+1$ st problem we wish to minimize

$$\|\Phi(\xi_{j+1}, \lambda_{j+1}, \sigma_{j+1})\|$$

(20)

If ξ_j, λ_j , and σ_j give the solution of Eqs. (18) and (19) for the j th set of parameter values σ_j , then

$$\Phi(\xi_{j+1}, \lambda_{j+1}, \sigma_{j+1}) = \Phi(\xi_j + \Delta \xi_j, \lambda_j + \Delta \lambda_j, \sigma_j + \Delta \sigma_j)$$

$$\approx \Phi(\xi_j, \lambda_j, \sigma_j) + \left. \frac{\partial \Phi}{\partial \xi} \right|_{\xi=\xi_j} \Delta \xi_j$$

$$+ \left. \frac{\partial \Phi}{\partial \lambda} \right|_{\lambda=\lambda_j} \Delta \lambda_j + \left. \frac{\partial \Phi}{\partial \sigma} \right|_{\sigma=\sigma_j} \Delta \sigma_j$$

Therefore, to find a reasonable approximation to ξ_{j+1}, λ_{j+1} from ξ_j, λ_j minimize

$$\left\| \Phi(\xi_j, \lambda_j, \sigma_j) + \left. \frac{\partial \Phi}{\partial \xi} \right|_{\xi=\xi_j} \Delta \xi_j + \left. \frac{\partial \Phi}{\partial \lambda} \right|_{\lambda=\lambda_j} \Delta \lambda_j + \left. \frac{\partial \Phi}{\partial \sigma} \right|_{\sigma=\sigma_j} \Delta \sigma_j \right\|$$

(21)

with respect to $\Delta \xi_j, \Delta \lambda_j$, so that

$$\xi_{j+1}^0 = \xi_j + \Delta \xi_j, \quad \lambda_{j+1}^0 = \lambda_j + \Delta \lambda_j$$

It is clear that the matrix $\partial \Phi / \partial \sigma$ has the form

$$\frac{\partial \Phi}{\partial \sigma} = \begin{bmatrix} BZ_J \\ -Z_1 \\ -Z_2 \\ \vdots \\ -Z_{J-1} \end{bmatrix}$$

where

$$Z_i = \frac{\partial \mathbf{u}_i}{\partial \sigma}(t_i), \quad i = 1, 2, \dots, J$$

4. Computations

The algorithms described have been implemented in a FORTRAN V subroutine for the UNIVAC 1108.¹ The solutions of the initial value problems are obtained using a variable order Adams–Moulton type integration.² The nonlinear systems (Expressions 8 and 19) are minimized by a modified Newton method,³ which makes use of singular value decompositions and pseudo-inverses to improve the global convergence characteristics of Newton's method. All linear systems are solved in a "least-squares" sense, using a singular value decomposition.

The subroutine¹ automatically selects the breakpoints for the multiple shots on the basis of the maximum component of the variational system. In addition, to insure the desired accuracy, the subroutine sometimes adds, moves, or deletes breakpoints, and corresponding initial values are found. Although this subroutine is presently limited to handling two-point problems with linear constraints, it is not difficult to extend the subroutine to handle multipoint problems with nonlinear constraints.

The work reported in this article was made possible by the many contributions of F. T. Krogh.

References

1. Keller, H. B., *Numerical Methods For Two-Point Boundary-Value Problems*. Blaisdell Publishing Company, Waltham, Mass., 1968.
2. Conte, S. D., "The Numerical Solution of Linear Boundary Value Problems," *SIAM Review* 8, pp. 309–321, 1966.

B. On the Design of a Primitive Fast Fourier Transformation Process, F. T. Krogh

1. Introduction

The basic fast Fourier transform (FFT; Ref. 1), is an efficient computational procedure for computing sums of the form

$$x_j = \sum_{k=0}^{N-1} c_k W^{jk}, \quad j = 0, 1, \dots, N-1 \quad (1)$$

where $W = e^{2\pi i/N}$, $i = (-1)^{1/2}$, the x_j and c_k are complex, and N is composite. It has been found that restricting N

to be of the form $N = 2^M$, M a positive integer, is not too severe a restriction in practice; this is assumed to be the case here. It is also assumed that the input array c is replaced by the output array x in the shared array of computer storage A , during the course of the calculation.

A strong case can be made (see e. g., Ref. 2) for using a procedure which carries out the calculations in Eq. (1) as the nucleus of a number of programs for computing Fourier transforms and associated tasks. These calculations may be carried out by specialized hardware, or by an efficiently programmed subroutine. Two features are described which give added flexibility to this primitive fast Fourier transform procedure with essentially no cost as far as complexity or efficiency are concerned. The efficiency of the radix 2 and radix $4+2$ FFT algorithms are compared, and an algorithm is given for the scrambling process which is required by both of these algorithms. The ideas in this article have been incorporated in the subroutine FFT,⁴ which in turn is called by the general-purpose subroutines.⁵⁻⁸ A recent bibliography on the FFT is given by Singleton (Ref. 3).

2. The Two Features

We suggest that the primitive FFT process permits the user to specify:

- (1) An index i_d which gives the distance in core between successive real parts and successive imaginary parts of the numbers in A .
- (2) The address a_r of the real part and the address a_i of the imaginary part of the first (complex) number in A .

Typically, the user is free to specify only a_r , with $a_i = a_r + 1$, and $i_d = 2$ by default. (This assumes that the compiler is one which stores the imaginary part of a complex number in the cell immediately following the real part.) The arbitrary value for i_d enables one to compute multidimensional transforms without the need to

⁴Krogh, F. T., *FFT—Primitive Fast Fourier Transform*, Oct. 1969 (JPL internal document).

⁵Krogh, F. T., *CFT—Multidimensional Complex Fourier Transform*, Oct. 1969 (JPL internal document).

⁶Krogh, F. T., *RFT—Multidimensional Real Fourier Transform*, Oct. 1969 (JPL internal document).

⁷Krogh, F. T., *RFT1—One-Dimensional Real Fourier Transform*, Oct. 1969 (JPL internal document).

⁸Krogh, F. T., *SCT—Multidimensional Sine, Cosine, and Sine-Cosine Transforms*, Oct. 1969 (JPL internal document).

¹Keener, J. P., *SPBVP/BVPQS*, Sept. 1969 (JPL internal document).

²Krogh, F. T., *VODQ/SVDQ/DVDQ—Variable Order Integrators for the Numerical Solution of Ordinary Differential Equations*, May 1969 (JPL internal document).

³Hanson, R. J., *NWTR/NWTR2*, Aug. 1969 (JPL internal document).

move the data into contiguous locations. In the case of multidimensional sine or cosine transforms, a_R and a_I must be set independently if the unnecessary moving of data is to be avoided. (The procedure used for computing sine and cosine transforms is given in Ref. 2. The obvious extension to the multidimensional case is given in Footnote 8.)

The independent specification of a_R and a_I also leads to a minor simplification in computing

$$c_k = \frac{1}{N} \sum_{j=0}^{N-1} x_j W^{-jk}, \quad k = 0, 1, \dots, N-1 \quad (2)$$

which is the inverse of the process in Eq. (1). Clearly,

$$c_k^* = \frac{1}{N} \sum_{j=0}^{N-1} x_j^* W^{jk}, \quad k = 0, 1, \dots, N-1 \quad (3)$$

and

$$(ic_k^*) = \frac{1}{N} \sum_{j=0}^{N-1} (ix_j^*) W^{jk}, \quad k = 0, 1, \dots, N-1 \quad (4)$$

where * is used to indicate a complex conjugate. The usual way to compute Eq. (2), using a process for computing Eq. (1), follows from Eq. (3): divide the input array by N , take its complex conjugate, carry out the same procedure as is used for Eq. (1), and then take the complex conjugate of the result. Since multiplying the conjugate of a complex number by i reverses the real and imaginary parts of the original number, it is clear from Eq. (4) that one could replace the above conjugations with an interchange of the real and imaginary parts. Such a procedure would of course be more clumsy. However, exactly the same effect can be produced by calling the procedure used for computing Eq. (1) with a_R and a_I interchanged. This is the minor simplification mentioned at the beginning of this paragraph.

3. Comparison of the Radix 2 and Radix 2 + 4 Algorithms

Before the FFT, the calculations indicated in Eq. (1) were usually carried out using nested multiplications and

additions to compute x_j as an N th degree polynomial in W^j , a process which requires $N-1$ operations for each j , where an operation is defined to be a complex multiplication and a complex addition. Ignoring the fact that these calculations can be simplified for $j = \ell N/8$, $\ell = 0, 1, \dots, 7$, this procedure requires a total of $N(N-1)$ operations.

The general idea for the radix 2 algorithm is obtained as follows. Let

$$u_j = \sum_{k=0}^{(N/2)-1} c_{2k} W^{2jk} \quad (5a)$$

and

$$v_j = \sum_{k=0}^{(N/2)-1} c_{2k+1} W^{2jk}, \quad j = 0, 1, \dots, \frac{1}{2}N-1 \quad (5b)$$

Since $W^{N/2} = -1$, it is clear that

$$\left. \begin{aligned} x_j &= u_j + v_j W^j \\ x_{(N/2)+j} &= u_j - v_j W^j \end{aligned} \right\} j = 0, 1, \dots, \frac{1}{2}N-1 \quad (6)$$

By treating $j=0$ ($W^0=1$) and $j=N/4$ ($W^{N/4}=i$, and multiplication by i does not actually require a multiplication) as special cases, the calculations indicated in Eq. (6) require

$$\left. \begin{aligned} &N \text{ complex additions} \\ \frac{1}{2}N-2, & \quad n \geq 4 \\ 0, & \quad n \leq 2 \end{aligned} \right\} \text{complex multiplications} \quad (7)$$

Using the old procedure, the calculations indicated in Eq. (5) require

$$2[\frac{1}{2}N(\frac{1}{2}N-1)]$$

operations. Adding this to the calculations required by Eq. (6), we find the number of operations has been reduced by nearly half. Of course, the same idea can be used to compute u_j and v_j more efficiently, etc. This process repeated M times ($N=2^M$) gives the radix 2 algorithm.

The radix 4 + 2 algorithm is obtained in a similar way. Let

$$\left. \begin{aligned} \alpha_j &= \sum_{k=0}^{(N/4)-1} c_{4k} W^{4jk}, & \beta_j &= \sum_{k=0}^{(N/4)-1} c_{4k+1} W^{4jk} \\ \gamma_j &= \sum_{k=0}^{(N/4)-1} c_{4k+2} W^{4jk}, & \delta_j &= \sum_{k=0}^{(N/4)-1} c_{4k+3} W^{4jk+3} \end{aligned} \right\} j = 0, 1, \dots, \frac{1}{4}N - 1 \quad (8)$$

It is easy to see that

$$\left. \begin{aligned} x_j &= \alpha_j + \beta_j W^j + \gamma_j W^{2j} + \delta_j W^{3j} \\ x_{j+(N/4)} &= \alpha_j + i\beta_j W^j - \gamma_j W^{2j} - i\delta_j W^{3j} \\ x_{j+(N/2)} &= \alpha_j - \beta_j W^j + \gamma_j W^{2j} - \delta_j W^{3j} \\ x_{j+(3N/4)} &= \alpha_j - i\beta_j W^j - \gamma_j W^{2j} + i\delta_j W^{3j} \end{aligned} \right\} j = 0, 1, \dots, \frac{1}{4}N - 1 \quad (9)$$

By treating $j = 0$ and $j = N/8$ [$W^{N/8} = \frac{1}{2}(2)^{1/2}(1 + i)$] as special cases, the calculations indicated in Eq. (9) require the equivalent of

$$\left. \begin{aligned} 8 \text{ complex additions} & & N &= 4 \\ 2N + 1 \text{ complex additions} & & & \\ \frac{3}{4}N - 5 \text{ complex multiplications} & \left\{ \begin{aligned} N &\geq 8 \end{aligned} \right. & (10) \end{aligned}$$

This process can then be applied in the same way to the calculations indicated in Eq. (8), etc. Repeated $M/2$ times when M is even, and $(M - 1)/2$ times with a radix 2 algorithm at the beginning when M is odd, this process gives the radix 4 + 2 algorithm.

A complete count of the real arithmetic operations required by these two algorithms is given in Table 1. (A complex multiplication is counted as 4 real multiplications and 2 real additions, and a complex addition is counted as 2 real additions.) Also given in Table 1 is the number of operations required for the complex Fourier transform of real data (Footnote 7), using both the radix 2 and the radix 4 + 2 transform in the primitive Fourier transform process, and the number of operations required, using the old technique mentioned at the beginning of this subsection. We believe the slightly greater complexity of the radix 4 + 2 algorithm is more than compensated for by its improved efficiency over the radix 2 algorithm. The radix 8 + 4 + 2 algorithm is just a little more efficient than the radix 4 + 2 algorithm, but not enough so in our opinion to justify the additional program complexity. (To a first approximation, when N is large the radix 8 + 4 + 2 algorithm requires 8/9 the number of real multiplications used by the radix 4 + 2 algorithm.)

4. An Algorithm for Scrambling the Array

It is well known that if the FFT is to be computed without using an extra array of storage, elements of the array with indices in reverse bit order must be interchanged either before or after the main computation, depending on the form of the FFT used. This is frequently accomplished in FORTRAN by using a nested sequence of DO loops, see e.g., Ref. 4. The algorithm given below has the following minor advantages over the nested DO loop method: it requires less program storage (at least this is true on the UNIVAC 1108), it is easier to convert the subroutine to have a different maximum number of points, and it is slightly more efficient, especially when the number of points is less than the maximum number permitted.

Let AR(1) and AI(1) denote the real and imaginary parts of the first element in the array, and let i_d denote the index which gives the distance in core between successive numbers in the array. Let $M = \log_2 N$, and let $K(L) = i_d \cdot 2^{M-L}$, $L = 1, 2, \dots, M$. The algorithm is given in flowchart form in Fig. 1. The basic idea is the same as the idea used for the nested DO loop method. The indices I and IR are in reverse bit order if one is subtracted from both and the results divided by i_d .

References

1. Cooley, J. W., and Tukey, J. W., "An Algorithm for the Machine Calculation of Complex Fourier Series," *Math. of Comp.*, Vol. 19, pp. 297-301, 1965.
2. Cooley, J. W., Lewis, P. A. W., and Welch, P. D., *The Fast Fourier Transform Algorithm: Programming and Accuracy Considerations in the Calculation of Sine, Cosine, and Laplace Transforms*. IBM Watson Research Center, Yorktown Heights, New York, 1969.

3. Singleton, R. C., "A Short Bibliography on the Fast Fourier Transform," *IEEE Trans. Audio Electroacoustics*, Vol. AU-17, pp. 166-169, 1969.

4. Bergland, G. D., "A Radix-Eight Fast Fourier Transform Subroutine for Real-Valued Series," *IEEE Trans. Audio Electroacoustics*, Vol. AU-17, pp. 138-144, 1969.

Table 1. Number of real operations required to compute N point Fourier transforms

M	N	Radix 4 + 2 algorithm				Radix 2 algorithm				Old Way
		$a_c(N)$	$m_c(N)$	$a_r(N)$	$m_r(N)$	$a_c(N)$	$m_c(N)$	$a_r(N)$	$m_r(N)$	$a_c(N) = m_c(N)$
1	2	4	0	2	0	4	0	2	0	8
2	4	16	0	8	0	16	0	8	0	48
3	8	52	4	30	4	52	8	30	4	224
4	16	144	28	86	16	148	40	86	20	960
5	32	376	92	218	56	388	136	222	68	3,968
6	64	920	284	530	152	964	392	542	196	16,128
7	128	2,200	732	1,234	408	2,308	1,032	1,278	516	65,024
8	256	5,080	1,884	2,834	984	5,380	2,568	2,942	1,284	261,120
9	512	11,608	4,444	6,354	2,392	12,292	6,152	6,654	3,076	1,046,528
10	1,024	25,944	10,588	14,162	5,464	27,652	14,344	14,846	7,172	4,190,208
11	2,048	57,688	23,900	31,058	12,632	61,444	32,776	32,776	16,388	16,769,024
12	4,096	126,296	54,620	67,922	27,992	135,172	73,736	71,678	36,868	67,092,480
13	8,192	275,800	120,156	146,770	62,808	294,916	163,848	155,646	81,924	268,402,688
14	16,384	595,288	267,612	316,754	136,536	638,980	360,456	335,870	180,228	1,073,676,288
		$a_c(N) = \text{number of real additions for an } N \text{ point complex Fourier transform} = \begin{cases} (33MN - 26N + 32)/12 & M \text{ even} \\ (33MN - 25N + 32)/12 & M \text{ odd} \end{cases} \left. \begin{matrix} \text{radix 4 + 2} \\ \text{radix 2} \\ \text{old way} \end{matrix} \right\}$								
		$m_c(N) = \text{number of real multiplications for an } N \text{ point complex Fourier transform} = \begin{cases} (9MN - 28N + 40)/6 & M \text{ even} \\ (9MN - 29N + 40)/6 & M \text{ odd} \end{cases} \left. \begin{matrix} \text{radix 4 + 2} \\ \text{radix 2} \\ \text{old way} \end{matrix} \right\}$								
		$a_r(N) = \text{number of real additions for an } N \text{ point real Fourier transform} = \begin{cases} \frac{1}{2}a_c(N) + 7(N-4)/6 & = (33MN + 2N - 80)/24 & M \text{ even} \\ \frac{1}{2}a_c(N) + (13N - 56)/12 & = (33MN + N - 80)/24 & M \text{ odd} \\ \frac{1}{2}a_c(N) + N - 4 & = (3MN - N - 4)/2 & \text{radix 2} \end{cases} \left. \begin{matrix} \text{radix 4 + 2} \\ \text{radix 2} \end{matrix} \right\}$								
		$m_r(N) = \text{number of real multiplications for an } N \text{ point real Fourier transform} = \begin{cases} \frac{1}{2}m_c(N) + (N-4)/6 & = (9MN - 26N + 32)/12 & M \text{ even} \\ \frac{1}{2}m_c(N) + (N-2)/3 & = (9MN - 25N + 32)/12 & M \text{ odd} \\ \frac{1}{2}m_c(N) & = MN - 3N + 4 & \text{radix 2} \end{cases} \left. \begin{matrix} \text{radix 4 + 2} \\ \text{radix 2} \end{matrix} \right\}$								

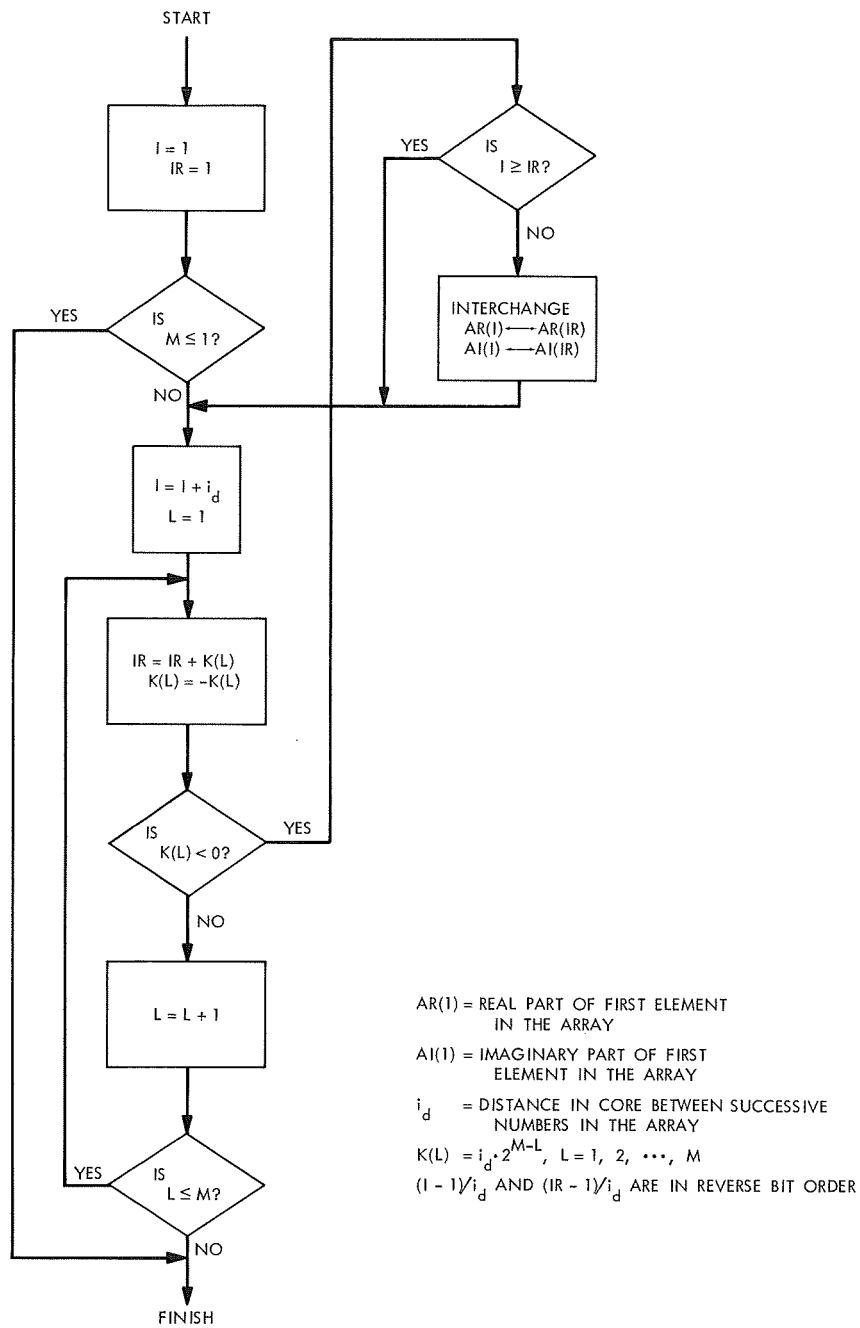


Fig. 1. Scrambling algorithm

III. Space Instruments

SPACE SCIENCES DIVISION

A. A Spherical Coil System for the Helium Vector Magnetometer, W. L. Momsen

1. Introduction

In the helium magnetometer a beam of near infrared light is modulated by the magnetic field applied to an absorption cell containing metastable helium atoms. Two magnetic field components must be generated by the coil system: a rotating sweep vector, and a dc feedback component that effectively nulls the ambient field. Both ac and dc fields are generated by the same coil system, which consists of three separate orthogonal pairs for triaxial operation.

In the *Mariner IV* magnetometer, the necessary fields were generated by a set of Helmholtz coils 6 in. in diameter. Since the coil power consumption is directly proportional to the square of the field generated, and to the fifth power of the coil radius, attempting to meet the design goal capability of measuring 10-G fields would have resulted in power consumption of 600 W. Therefore, a radical reduction in coil diameter was necessary. The absorption cell was reconfigured from a 1.5×2 -in. cylinder to a 1.25-in. sphere. The most efficient coils would lie as close to the cell surface as possible, since

power consumption increases rapidly with coil-cell distance. The present coils are made in the form of hemispherical shells to minimize this dimension.

In the following formulas, the effects of field inhomogeneity over the volume of the cell are neglected, and the calculations apply only for the field at the center of the coil system.

2. Computation of Coil Power Consumption and Weight¹

It can be shown that the field at the center of the coil is

$$B = \frac{\mu_0}{4} n i \frac{(2\theta_2 - 2\theta_1 + \sin 2\theta_1 - \sin 2\theta_2)}{(R_2 + R_1)(\theta_2 - \theta_1)} \quad (1)$$

where n is the number of turns, and θ_1 , θ_2 , R_1 , R_2 are defined in Fig. 1. Solving for current i

$$i = \frac{4}{\mu_0} \frac{B}{n} \frac{(R_2 + R_1)(\theta_2 - \theta_1)}{(2\theta_2 - 2\theta_1 - \sin 2\theta_2 + \sin 2\theta_1)} \quad (2)$$

in MKS units.

¹All angles are in radians.

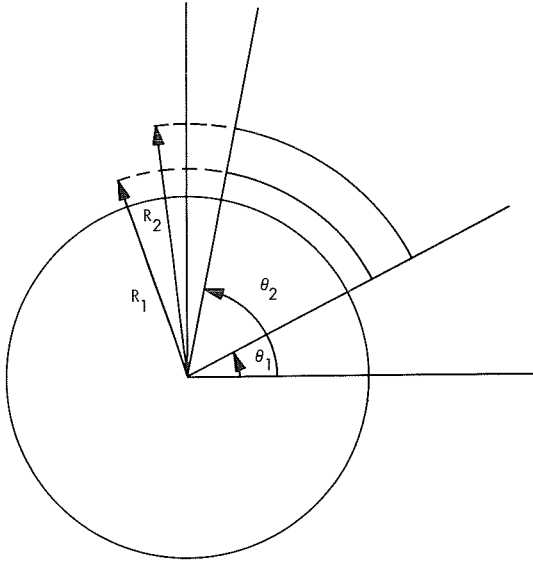


Fig. 1. Coil geometry

For R in inches and B in gauss, Eq. (2) becomes

$$i = 8.06 \frac{B}{n} \frac{(R_2 + R_1)(\theta_2 - \theta_1)}{(2\theta_2 - 2\theta_1 - \sin 2\theta_2 + \sin 2\theta_1)} \quad (3)$$

In terms of wire constants K_1 (turns/in.²), K_2 (ohms/ thousand feet), and K_3 (lb/thousand feet), dc coil resistance Ω and weight W are:

$$\Omega = \frac{2\pi}{36,000} K_1 K_2 (R_2^3 - R_1^3) (\cos \theta_1 - \cos \theta_2) \quad (4)$$

$$W = \frac{K_3}{K_1} \Omega \quad (5)$$

Substituting Eqs. (3) and (4) directly into the relationship $P = i^2 \Omega$ results in a complicated expression for power; Eqs. (3) and (4) are more conveniently solved independently for i and Ω , and the results used to calculate coil power.

From known coil geometry (R_1 , R_2 , θ_1 , θ_2) and desired field strength, Eqs. (3), (4), and (5) can be used to determine dc coil power, weight, and resistance for any given wire size. Note that the equations are for a single hemispherical coil. For a set of two coils in series B , Ω , and W are doubled, although i remains unchanged.

Power and weight are essentially independent of wire size, although a slight increase in efficiency can be realized by using smaller wire sizes. Impedance match-

ing between the coil and driver can be accomplished by adjusting wire size to change dc coil resistance.

3. Technique for Winding Hemispherical Shell Coils

Wire cannot easily be wound directly on the surface of a sphere. Winding upward can be accomplished, since each successive turn is supported by the previous turn. However, attempting to wind the spherical surface back down over the first layer results in the wire slipping down the slope. Although each turn could be bonded as it is being wound, the following method proved easier to implement.

The spherical shell geometry is translated into a shape easier to wind and an aluminum coil form fabricated. The coil is wound on the intermediate form, using enameled wire coated with an adhesive that softens with the application of heat or a suitable solvent. As it is wound on the form, the solvent is applied to bond the turns. The coil is then removed from the form and placed between mating hemispherical dies under moderate pressure and heat. As the adhesive softens, the coil assumes the shape of the dies. Upon cooling it is removed and will conform to the spherical surface.

For triaxial operation, two coils are made for each axis, and the complete coil system consists of three nesting coil pairs.

Referring to Fig. 2, the cross section of the coil can be conveniently divided into three regions. The vertical

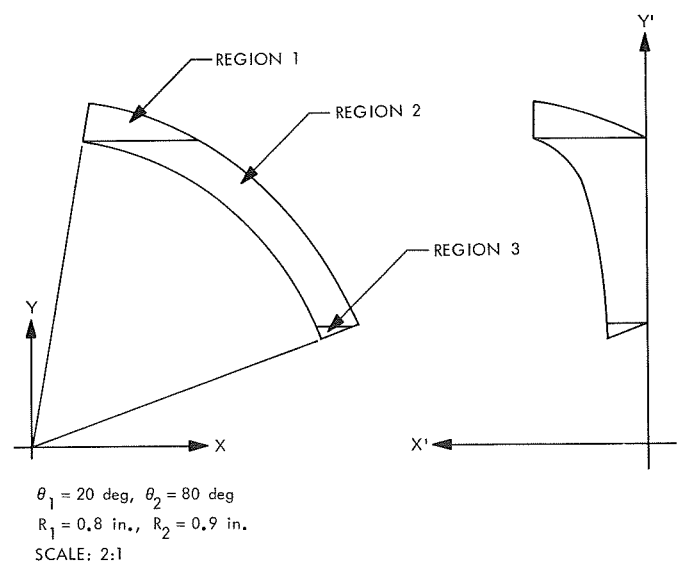


Fig. 2. Transformation of coordinates

edge of the coil form corresponds to the outer surface of the finished coil.

The width of the form along the X' axis for any given vertical coordinate Y' is the corresponding width of the coil along the X axis for the same vertical coordinate ($Y = Y'$). For the three regions X' can be expressed as a function of Y by the following set of equations:

$$\begin{aligned} X'_1 &= R_2 \cos \theta_1 - [R_1^2 - (R_2 \sin \theta_1)^2]^{1/2} \\ &\quad - Y \cot \theta_1 + [R_1^2 - Y^2]^{1/2} \\ R_1 \sin \theta_1 &< Y < R_2 \sin \theta_1 \end{aligned} \quad (6)$$

$$\begin{aligned} X'_2 &= [R_2^2 - Y^2]^{1/2} - [R_1^2 - Y^2]^{1/2} \\ R_2 \sin \theta_1 &< Y < R_1 \sin \theta_2 \end{aligned} \quad (7)$$

$$\begin{aligned} X'_3 &= [R_2^2 - (R_1 \sin \theta_2)^2]^{1/2} - R_1 \cos \theta_2 \\ &\quad - [R_2^2 - Y^2]^{1/2} + Y \cot \theta_2 \\ R_1 \sin \theta_2 &< Y < R_2 \sin \theta_2 \end{aligned} \quad (8)$$

Depending on specific coil geometry, the curves bounding regions 1 and 3 may, in some cases, be replaced by straight line segments.

The coil forms were fabricated on a lathe, using computer-generated $X'-Y'$ pairs as handwheel settings. Y , incremented by some convenient value, was substituted into Eqs. (6), (7), and (8), which were solved for X' .

The forms were made in two parts, as shown in Fig. 3, held together on a threaded mandrel during winding, and disassembled for coil removal. The final lays of the winding were wound to conform to a template (shown in the upper portion of Fig. 3) in accordance with the equation defining the upper bound of region 3.

The molds used for final shaping of the coils consisted of two concentric hemispheres with radii corresponding to the inner and outer radii of the desired coil. The coil was placed between the two sides of the mold, and assembled (with an inner cylindrical spacer and outer ring to establish the upper and lower coil angles) on a threaded rod. A spring was placed on the rod to apply pressure to the assembly, which was then heated to soften the adhesive. The spring pressure then forced the wire to conform to the shape of the mold.

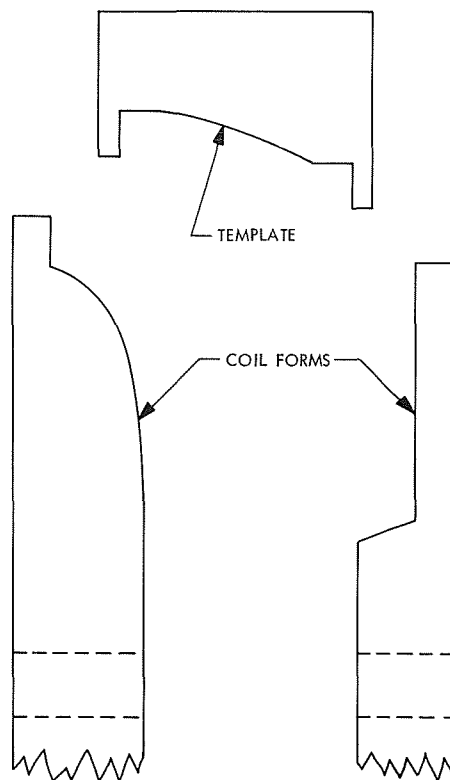


Fig. 3. Coil forms and template

4. Conclusions

Figure 4 is a plot of coil power versus weight for triaxial coil sets of various inside diameters. The most efficient operating point for each set corresponds to the minimum on the associated curve. Power consumption decreases with increasing coil weight only until the coil diameter becomes so large that each turn added, being at a greater distance from the center, reduces rather than increases coil efficiency. To take advantage of the minimum point on the curve, the inside coil diameter must be relatively small.

Since the field generated by a single turn contributes to a vector sum at the center of the coil system, the most efficient part of the coil lies directly above the center. Coil efficiency can be kept high by using values of θ_1 and θ_2 as large as possible, consistent with other requirements, e.g., the need for apertures in the coils.

The parameters describing the triaxial coil set that was constructed are listed in Table 1.

In the course of tests performed to evaluate various coil configurations for the *Pioneer F* and *G* Magnetometer

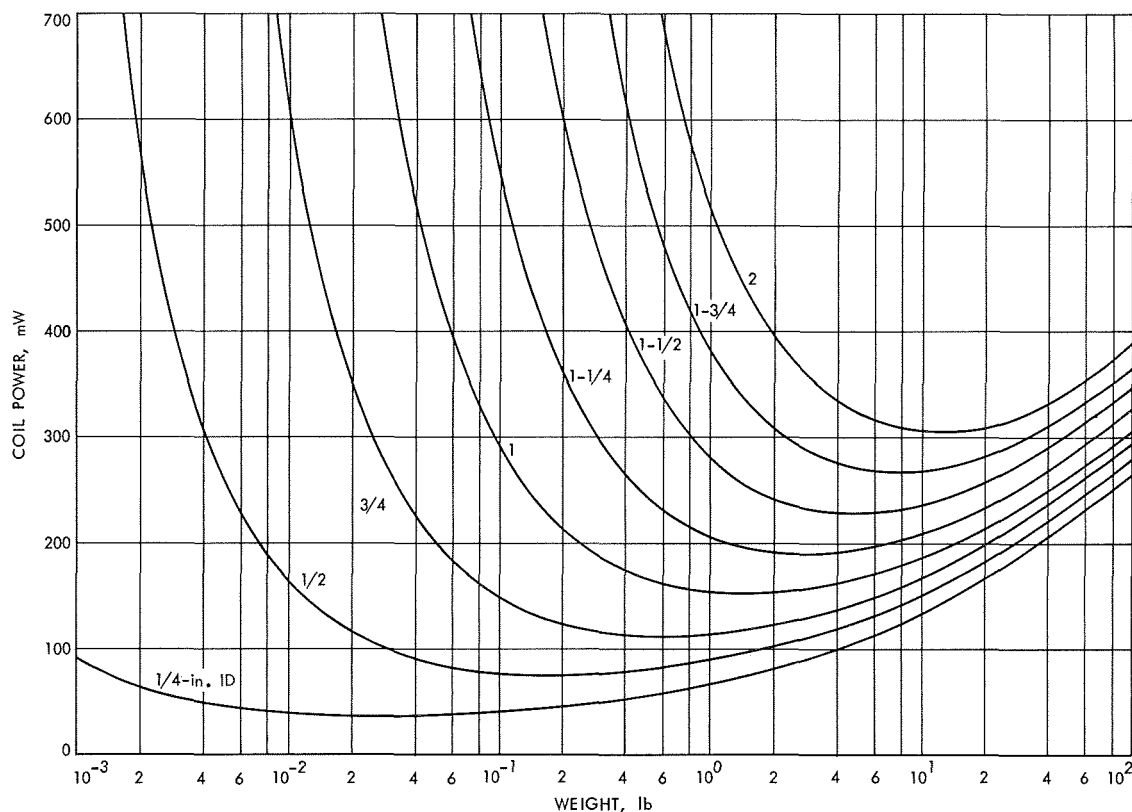


Fig. 4. Power versus weight for spherical coil diameters

Program, a large apparent offset in the field measurement was observed during high field operation. This was ascribed to inhomogeneity of the field within the absorption cell due to the coils. The more efficient coils with high field capability were therefore abandoned in favor of a Helmholtz set to improve field linearity. Further testing revealed, however, that field homogeneity was not the source of the offset; the more efficient coils merely allowed operation in high fields where this phenomenon became apparent.

The helium magnetometer is a closed loop self-nulling system. The error signal is derived from the intensity of

the light passing through the cell, and consists of an ac signal of the same frequency as the sweep field (SPS 37-30, Vol. IV, p. 131). The sweep field also amplitude modulates the plasma in the cell to produce a signal of the same frequency, but independent of the ambient magnetic field.

This effect is observed only for sweep fields greater than 1 G, appearing in the data as an offset proportional to the sweep amplitude. Several approaches to eliminating this problem are currently under investigation. Upon correction of the problem, a high field helium magnetometer can be constructed using the coils described.

Table 1. Prototype triaxial coil parameters

Coil set	Turns	R_1	R_2	θ_1	θ_2	Power, watts	Resistance, ohms	Weight, lb
Inner	884	0.675	0.775	21.5	82.5	0.077	100	0.096
Center	841	0.775	0.875	20.5	71.5	0.149	100	0.097
Outer	943	0.875	0.975	20.5	71.5	0.187	126	0.121
								Total 0.314

All coils are wound with No. 32 single enamel, single-coated Bondeze wire. Power is for a 10-G dc field at the center of the coil system for each of the three axes, respectively. θ_1 and θ_2 were chosen to provide apertures for the light beam and cell tip-off.

IV. Lunar and Planetary Sciences

SPACE SCIENCES DIVISION

A. Effect of Fe, Ca, and Mn Concentrations on Thermoluminescence Intensity of Meteoritic Enstatites, R. T. Greer¹

Enstatite grains (approximately 150–180 μm) were separated from 13 different chondritic and achondritic meteorite specimens, and relationships between total thermoluminescence output and elemental composition for the single phase were sought. There was sufficient variability in element concentrations exhibited by the specimens to permit an evaluation of the influence of composition on thermoluminescence emission. Seven of the eleven enstatite chondrites (Adhi Kot, Abee, Atlanta, Hvittis, Khairpur, Jajh deh Kot Lalu, and Blithfield) and six of the nine enstatite achondrites (Bishopville, Cumberland Falls, Khor Temiki, Norton County, Pesyanoe, and Shallowater) which have been identified as meteorites are included in this work.

The investigation was restricted to the pyroxene, enstatite. All pyroxene structures link SiO_4 tetrahedra by sharing two of the four corners, forming continuous $(\text{SiO}_3)_n$ chains. The 5.3-Å repeat distance along the length of the chain defines the c parameter of the unit cell. The cations (Ca, Fe, Mg, etc.) link the chains laterally. In

enstatite (MgSiO_3), Mg or (Mg, Fe) atoms laterally link the $(\text{SiO}_3)_n$ chains. Ions other than Mg and Fe can be present in orthopyroxenes and commonly include Ca and Mn; usually the sum of these constituents is less than 10 mol %. The polymorphs with the structure of enstatite contain less than 30 mol % FeSiO_3 and less than 15 mol % of CaSiO_3 , as defined by the structural grouping for pyroxenes in the field $\text{CaMgSi}_2\text{O}_6$ – $\text{CaFeSi}_2\text{O}_6$ – MgSiO_3 – FeSiO_3 . The majority of pyroxenes can be considered to be members of this four-component system. In both the enstatite chondrite and achondrite specimens, many of the metals which commonly substitute for the Mg cation in the pyroxene structure are present in low concentrations, or below detection limits. Therefore, the disposition of the Fe, Mn, and Ca cations are of primary interest.

The enstatite subspecimens were examined by an optical microscope, X-ray diffraction, an electron microprobe, and a thermoluminescence detection unit. A Perkin-Elmer differential scanning calorimeter standard ¼-in.-diameter sample holder and unit with a constant rate heating element, an RCA 6199 photomultiplier tube for detecting the natural thermoluminescence emission, a heating rate of 80°C/min, a heating range of 25–500°C, and a dry N_2 gas flow of 20 cm^3/min were chosen for the experiment.

¹NRC–NASA Resident Research Associate.

Figure 1 represents the variation of integrated thermoluminescence emittance for enstatite from individual meteorite specimens with respect to Fe concentration. A number of significant relationships are featured on this diagram. The trend of the positions of the specimens on the diagram, considered from the lower right corner to the upper left corner, is in agreement with the following relationships: change of meteorite class, decreasing Fe in the pyroxene structure, and increase in thermoluminescence output. The theoretical expression to fit the data (using the least-squares principle) is

$$y = -2.601 - 0.866 x$$

Thus, the total thermoluminescence output is inversely proportional to the Fe concentration.

Multivariate statistical analytical techniques were utilized to process the data so as to evaluate the significance of the relationships that are indicated by the measurements. A significant negative correlation at the 0.1% confidence level is evident for the total thermoluminescence output and Fe. However, no significant correlations at the 10% level for total thermoluminescence output-Mn, total thermoluminescence output-Ca, or total thermoluminescence output-age (cosmic ray exposure or natural radioactivity) were found. The partial sample correlation coefficient $r_{\text{thermoluminescence, Fe/age}}$, which expresses the correlation of total thermoluminescence output and Fe independent of (rather than ignoring) the effect of age, is both negative and significant at 0.1%, confirming an observed and presumed casual relationship between these variables.

Orthopyroxene is second only to olivine in abundance as a silicate in meteorites. It can be further classified

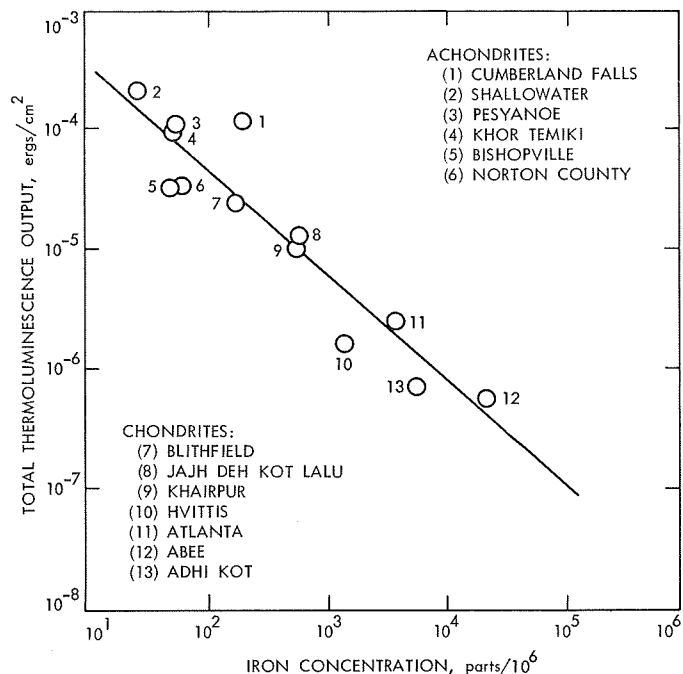


Fig. 1. Thermoluminescence response as related to Fe concentration for enstatite

according to its iron content as enstatite, less than 10 mol % FeSiO_3 ; bronzite, 10–20 mol % FeSiO_3 and hypersthene, 20–30 mol % FeSiO_3 . Thus, a prediction can be made on the basis of the linear relations shown in Fig. 1 that bronzite would exhibit a lower thermoluminescence energy density than enstatite, and that hypersthene would exhibit the lowest thermoluminescence emission in the series. Other relatively high Fe pyroxenes, such as pigeonite, would also be expected to exhibit extremely low or no detectable thermoluminescence emission.

V. Physics

SPACE SCIENCES DIVISION

A. The de Haas-van Alphen Effect in Indium,

R. W. Vaughan and D. D. Elleman

1. Introduction

Measurement of the de Haas-van Alphen effect is one of the most effective techniques to examine the Fermi surface in metallic and semimetallic materials. The experimental apparatus, a field-modulated magnetometer, uses a frequency modulation method to measure a small component of the magnetic susceptibility which oscillates as the value of an externally applied magnetic field changes. The amplitude and frequency of these susceptibility oscillations furnish direct measurements of the geometry of the Fermi surface, and by collecting data as a function of crystal orientation, temperature, and magnetic field, it is possible to characterize the Fermi surface in detail.

In studying materials where several cross-sections of the Fermi surface contribute simultaneously to the de Haas-van Alphen effect, problems are encountered in attempting to accurately measure the amplitudes and frequencies of the various superimposed oscillations. To overcome these problems and further improve the sensitivity of the apparatus, we have developed a scheme involving the use of numerical Fourier techniques to

invert the initial data and thus accurately determine both the frequencies and amplitudes of such multiple oscillations (Ref. 1).

2. Discussion

The Fermi surface and electronic structure of indium have been extensively investigated in the past few years (Refs. 2-5). Most of these studies, however, have dealt with only the third-zone electron surface, and little is known about the second-zone hole surface which comprises the largest portion of the indium Fermi surface. The only reported de Haas-van Alphen data on this second zone surface is contained in a letter by O'Sullivan and Schirber (Ref. 4). These authors report the value of the second-zone de Haas-van Alphen frequency at five symmetry points, and indicated that the amplitude of the frequencies behaved strangely, exhibiting a complex beating pattern which prevented them from determining any effective mass values.

We have examined the de Haas-van Alphen effect in indium using the Fourier transform technique discussed above and have been able to explain these reported abnormalities. Figure 1 is a power spectrum obtained by Fourier transforming an experimental de Haas-van Alphen

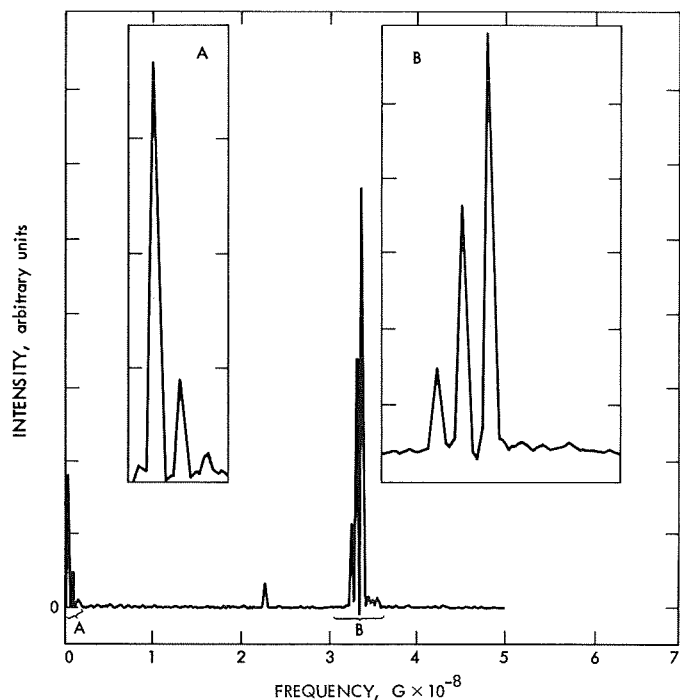


Fig. 1. Power spectrum of de Haas-van Alphen data for Indium at 1.23°K and magnetic field between 37.5 and 36.2 kG and near the [111] crystallographic axis

record for an indium crystal where the magnetic field is directed nearly parallel (within 1 deg) to the [111] crystallographic axis.

Two critical sections of this spectrum have been enlarged and shown in greater detail to demonstrate that the beating observed in the high-frequency data (second-zone hole surface) is due to the presence of several high-frequency components, and that these high frequencies differ in frequency by precisely the value of the low-frequency (third-zone) component. The fact that one expects to see only a single high-frequency oscillation due to the second-zone hole surface at this orientation, and that all the high frequencies differ by multiples of the low frequency, suggests that the spectrum observed is due to a single high-frequency term in the susceptibility interacting in a nonlinear fashion with the low-frequency component to the susceptibility and thus creating a number of combinational frequencies.

Data taken at other orientations in both the (100) and (011) planes gave similar results; in all cases the resultant complex spectrum could be accounted for in terms of combinational frequencies of the expected low and high-frequency de Haas-van Alphen oscillations. However, rapid changes in the spectrum occurred as the angles

changed, with rotations of only a few degrees being capable of changing the intensity of particular combinational frequencies by an order of magnitude.

Both the effect of temperature and the magnetic field on the amplitudes of the high-frequency oscillations were examined to identify the interaction responsible for the generation of such intense combinational frequencies.

Theory (Ref. 6) predicts that the effect of temperature T on the amplitude of a de Haas-van Alphen oscillation A will be to give a linear plot of $\ln(A/T)$ versus T , with the slope of the line being proportional to the effective mass of those sections of the Fermi surface contributing to that oscillation. In this case the spectrum consisted on a single intense peak with two shoulder peaks.

The values of the effective mass for the oscillations at this orientation were determined from a number of plots (Fig. 2), but over different ranges of magnetic field. The

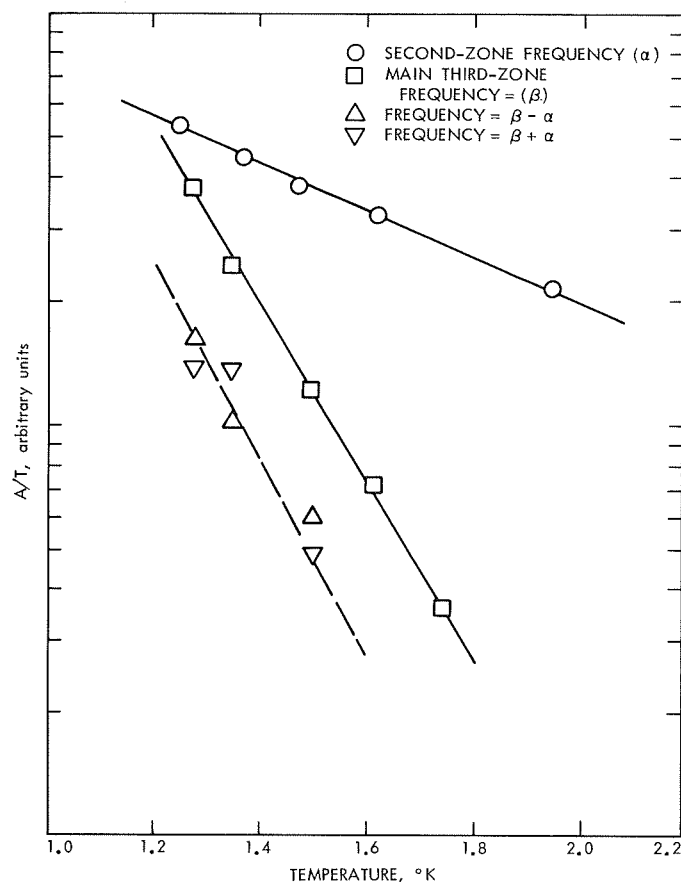


Fig. 2. Effective mass values for indium with [111] orientation (sets of data shifted vertically for clarity of presentation)

Table 1. Effective mass values for magnetic field parallel to the [111] axis

Fermi surface	Frequency	μ_{DHVA}	μ_{CR}^a
Third-zone electron surface	α	0.25 ± 0.005	0.25
Second-zone electron hole surface	β	1.35 ± 0.01	1.60
	$\beta - \alpha$	1.4 ± 0.2	
	$\beta + \alpha$	1.5 ± 0.1	

^aFrom Ref. 5.

resulting effective mass values are reported in Table 1. The error limits given in Table 1 are an indication of experimental scatter of data points and are large for the lower intensity shoulder peaks, mainly because the only amplitudes that could be measured were the highest fields and lowest temperatures.

Also reproduced in Table 1 are the effective mass values obtained from the cyclotron resonance experiments (Ref. 5). The low-frequency third-zone effective masses agree, but the high-frequency second-zone one differs by considerably more than the experimental error of either experimental technique. As the effective masses determined by these two techniques in similar metals, such as lead, have been found to agree well (Ref. 7), this represents an additional discrepancy to be accounted for.

Presently additional work is in progress to characterize this strong nonlinear interaction in terms of possible mechanism such as the Schoenberg B-H effect (Ref. 8), skin-depth modulation or interaction between de Haas-van Alphen and de Haas-Shubnikov effects.

References

1. Vaughan, R. W., Elleman, D. D., and McDonald, D. G., "The de Haas-van Alphen Effect in White Tin," *J. Phys. Chem. Solids*, 1969.
2. Ashcroft, N. W., and Lawrence, W. E., "Fermi Surface and Electronic Structure of Indium," *Phys. Rev.*, Vol. 175, pp. 938-955, 1968.
3. O'Sullivan, W. J., Schirber, J. E., and Anderson, J. R., "De Haas-van Alphen Study of the 2nd Zone Hole Surface of In," *Phys. Lett.*, Vol. 27A, pp. 144-145, 1968.
4. Hughes, A. J., and Shepherd, J. P. G., "De Haas-van Alphen Measurements and the Band Structure of Indium," *J. Phys.*, Sec. C, Vol. 2, pp. 661-671, 1969.
5. Mina, R. T., and Khaikin, M. S., "Investigation of the Fermi Surface and of Carrier Velocities in Indium by the Cyclotron Resonance Method," *Sov. Phys.—JETP*, Vol. 24, pp. 42-58, 1966; "Investigation of the Fermi Surface of Indium," *Sov. Phys.—JETP*, Vol. 21, pp. 75-82, 1965.

6. Lifshitz, I. M., and Kosevich, A. M., *Sov. Phys.—JETP*, Vol. 2, p. 636, 1956.
7. Phillips, R. A., and Gold, A. V., "Landau-Level Widths, Effective Masses, and Magnetic-Interaction Effects in Lead," *Phys. Rev.*, Vol. 178, pp. 932-948, 1969.
8. Shoenberg, D., "The Magnetic Interaction Effect, I. Interaction of Two Frequencies," *Canad. J. Phys.*, Vol. 46, pp. 1915-1934, 1968.

B. Integral Theorems in Dyadic Notation, H. D. Wahlquist

1. Introduction

Recent developments using the dyadic formalism of general relativity have required the formulation of the integral theorems involving comoving line, surface, and volume integrals for an arbitrarily moving frame of reference in a general spacetime manifold. This article provides a brief derivation and discussion of these theorems. Previous expositions of the differential analysis in the dyadic formalism, including derivations and applications, can be found in Refs. 1-2; SPS 37-27, Vol. IV, p. 139; and 37-38, Vol. IV, p. 159.

We write the space-time metric form as

$$ds^2 = -\frac{1}{\phi^2} [dt - A_\alpha dx^\alpha]^2 + h_{\alpha\beta} dx^\alpha dx^\beta, \quad \alpha, \beta = 1, 2, 3 \quad (1)$$

where the lines (t varies, $x^\alpha = \text{constant}$) are the world lines of an arbitrarily moving reference frame having local absolute acceleration \mathbf{a} , angular velocity $\boldsymbol{\Omega}$, and rate-of-strain \mathbf{S} . The comoving quotient space orthogonal to the timelike congruence has the positive definite metric form

$$dl^2 = h_{\alpha\beta} dx^\alpha dx^\beta \quad (2)$$

The metric coefficients A_α constitute the covariant components of a 3-vector \mathbf{A} in the quotient space, and together with the scalar coefficient ϕ serve as potentials for \mathbf{a} and $\boldsymbol{\Omega}$ which can be obtained from the potential equations

$$\phi \mathbf{a} = -\nabla \phi + \dot{\mathbf{A}} + \mathbf{S}^* \cdot \mathbf{A} \quad (3)$$

$$2\phi \boldsymbol{\Omega} = \nabla \times \mathbf{A} \quad (4)$$

where

$$\mathbf{S}^* \equiv \mathbf{S} - (\boldsymbol{\Omega} - \boldsymbol{\omega}) \times \mathbf{I} \quad (5)$$

The dot denotes proper time differentiation with respect to an orthonormal triad of basis vectors in the quotient space, rotating with angular velocity ω relative to Fermi propagated axes, and ∇ is the vector operator of covariant spatial differentiation in the quotient space. These operators do not commute; their commutation rules, and their relation to partial derivatives with respect to a coordinate system, are given in Ref. (2).

The potentials have additional significance as derivatives of the time coordinate; viz.,

$$\left. \begin{aligned} \phi &= \dot{t} = \frac{dt}{d\tau} \\ \mathbf{A} &= \nabla t \end{aligned} \right\} \quad (6)$$

so that by adopting proper time τ along the timelike congruence for the time coordinate, we may always achieve $\phi = 1$ with \mathbf{A} left unspecified. If the reference frame is not rotating ($\Omega = 0$), we may take $\mathbf{A} = 0$ with ϕ left unspecified. Finally, if both \mathbf{a} and Ω vanish, we may take $\phi = 1$ and $\mathbf{A} = 0$ (Gaussian coordinates). The integral theorems derived below, of course, take much simpler forms in these special cases.

2. Tetrad Representation of the Differential Forms

Let C be a space-time curve lying in the hypersurface $t = (\text{constant})$. The orthonormal tetrad components of an infinitesimal displacement along C can be written

$$d\tau^r = \left\{ -\frac{1}{\phi} \mathbf{A} \cdot d\mathbf{r}, d\mathbf{r} \right\} \quad (7)$$

where the component in the direction of the timelike congruence is given first, and the remaining three components are comprised in the 3-vector $d\mathbf{r} = \{dx^a\}$ which is a displacement vector on the projection of C into the quotient space. Writing the tetrad components of a 4-vector as

$$V_r = \{V_0, \mathbf{V}\} \quad (8)$$

we have the invariant differential form

$$V_r d\tau^r = \left[\mathbf{V} - \frac{V_0}{\phi} \mathbf{A} \right] \cdot d\mathbf{r} \quad (9)$$

Let Σ be a spacelike 2-surface, defined by the equations $t = (\text{constant})$ and $\psi(\chi^a) = (\text{constant})$, where the χ^a are co-

moving coordinates ($\dot{\psi} = 0$). The tensor extension (Ref. 3) for an element of Σ is

$$d\tau^{pq} = \frac{\delta \epsilon^{pqrs}}{\phi |\nabla \psi|} t_{,r} \psi_{,s} d\sigma \quad (10)$$

where $\delta = \pm 1$ as required to get the correct orientation and $d\sigma$ is the absolute area of a surface element on $\psi = (\text{constant})$ in the quotient space. That is, $d\sigma$ is defined in the usual way as the area ($dl_1 \cdot dl_2$) of a rectangular cell with sides of proper length dl_1 and dl_2 , where dl is defined in Eq. (2) with respect to the quotient metric $h_{\alpha\beta}$. We define the dual of the antisymmetric part of an arbitrary second rank 4-tensor F_{pq} by

$$H^{rs} \equiv \epsilon^{pqrs} F_{pq} \quad (11)$$

and represent H^{rs} by two 3-vectors

$$(\mathbf{H})^a \equiv H^{a0} = -H^{0a} \quad (12)$$

$$(\mathbf{E})_a \equiv \frac{1}{2} \epsilon_{abc} H^{bc} \quad (13)$$

$$H^{ab} = \epsilon^{abc} E_c \quad [a, b, c = 1, 2, 3] \quad (14)$$

Relating the 3 and 4-dimensional ϵ symbols by $\epsilon^{0abc} = \epsilon^{abc}$ and appropriate permutations, we can then write

$$\begin{aligned} F_{pq} d\tau^{pq} &= \frac{\delta}{\phi |\nabla \psi|} H^{rs} t_{,r} \psi_{,s} d\sigma \\ &= \delta \left[\mathbf{H} + \frac{1}{\phi} \mathbf{A} \times \mathbf{E} \right] \cdot d\sigma \end{aligned} \quad (15)$$

where

$$d\sigma \equiv \frac{\nabla \psi}{|\nabla \psi|} d\sigma \quad (16)$$

is the normal 3-vector representation of a surface element in $\psi = (\text{constant})$.

Let R be a spacelike hypersurface defined by $t = (\text{constant})$. The tensor extension of an element of R is

$$d\tau^{pqr} = \frac{\delta}{\phi} \epsilon^{pqr s} t_{,s} dv \quad (17)$$

where $\delta = \pm 1$ for orientation and dv is the absolute 3-volume of a volume element of the projection of R into the quotient space; i.e., again defined relative to the $h_{\alpha\beta}$

metric. We define the 4-vector J^s , dual to the completely antisymmetric part of a tensor K_{pqr} , by

$$J^s \equiv \epsilon^{pqrs} K_{pqr} \quad (18)$$

and represent its tetrad components as

$$J^s = \{J^0, \mathbf{J}\} \quad (19)$$

Then

$$\begin{aligned} K_{pqr} d\tau^{pqr} &= \frac{\delta}{\phi} J^s t_{,s} dv \\ &= \delta \left[J^0 + \frac{1}{\phi} \mathbf{A} \cdot \mathbf{J} \right] dv \end{aligned} \quad (20)$$

3. The Integral Theorems

Applying these equations to the case when $F_{pq} = V_{p;q}$, we find that the 3-vectors \mathbf{H} and \mathbf{E} are given by

$$\mathbf{H} = \nabla \times \mathbf{V} - 2V_0 \boldsymbol{\Omega} \quad (21)$$

$$\mathbf{E} = \nabla V_0 + V_0 \mathbf{a} - \dot{\mathbf{V}} - \mathbf{s}^* \cdot \mathbf{V} \quad (22)$$

and so the Stokes theorem

$$\oint_C V_r d\tau^r = \int_{\Sigma_C} V_{r;s} d\tau^{rs} \quad (23)$$

for the line integral of V_r around a closed curve C bounding a finite, simply connected region Σ_C of Σ has the dyadic expression

$$\oint_C \left[\mathbf{V} - \frac{V_0}{\phi} \mathbf{A} \right] \cdot d\mathbf{r} = \int_{\Sigma_C} \left\{ \nabla \times \mathbf{V} - 2V_0 \boldsymbol{\Omega} + \frac{1}{\phi} \mathbf{A} \times [\nabla V_0 + V_0 \mathbf{a} - \dot{\mathbf{V}} - \mathbf{s}^* \cdot \mathbf{V}] \right\} \cdot d\boldsymbol{\sigma} \quad (24)$$

where $d\boldsymbol{\sigma}$ is oriented such that on C ($d\boldsymbol{\sigma} \times d\mathbf{r}$) points to the interior of Σ_C . Since we can define an arbitrary 4-vector by giving any 3-vector \mathbf{V} and any 3-scalar V_0 , these are separately arbitrary. For $V_0 = 0$, the theorem reduces to

$$\oint_C \mathbf{V} \cdot d\mathbf{r} = \int_{\Sigma_C} \left[\nabla \times \mathbf{V} - \frac{1}{\phi} \mathbf{A} \times (\mathbf{V} + \mathbf{s}^* \cdot \mathbf{V}) \right] \cdot d\boldsymbol{\sigma} \quad (25)$$

but this result is actually equivalent to the general theorem. For if in Eq. (25) we put

$$\mathbf{V} \rightarrow \mathbf{V} - \frac{V_0}{\phi} \mathbf{A}$$

and then eliminate the derivatives of ϕ and \mathbf{A} on the right by means of the potential Eqs. (3) and (4) for \mathbf{a} and $\boldsymbol{\Omega}$, we recover Eq. (24). Equation (25) then can be taken as the general dyadic expression for Stokes' theorem.

Now letting $K_{pqr} = F_{pq;r}$ we have

$$J^s = \epsilon^{pqrs} F_{pq;r} = H^{rs}{}_{;r} \quad (26)$$

or

$$J^0 = \nabla \cdot \mathbf{H} + 2\boldsymbol{\Omega} \cdot \mathbf{E} \quad (27)$$

$$\mathbf{J} = -[(\nabla + \mathbf{a}) \times \mathbf{E} + \dot{\mathbf{H}} + (Tr\mathbf{S})\mathbf{H} - \mathbf{H} \cdot \mathbf{s}^*] \quad (28)$$

and so the Stokes (Gauss) theorem

$$\oint_{\Sigma} F_{rs} d\tau^{rs} = \int_{R_{\Sigma}} F_{rs;p} d\tau^{rsp} \quad (29)$$

for the integral of a tensor over a closed 2-surface Σ bounding a finite, simply-connected region R_{Σ} of R has the dyadic expression

$$\oint_{\Sigma} \left[\mathbf{H} + \frac{1}{\phi} \mathbf{A} \times \mathbf{E} \right] \cdot d\boldsymbol{\sigma} = \int_{R_{\Sigma}} \left\{ \nabla \cdot \mathbf{H} + 2\boldsymbol{\Omega} \cdot \mathbf{E} - \frac{1}{\phi} \mathbf{A} \cdot [(\nabla + \mathbf{a}) \times \mathbf{E} + \dot{\mathbf{H}} + (Tr\mathbf{S})\mathbf{H} - \mathbf{H} \cdot \mathbf{s}^*] \right\} dv \quad (30)$$

where $\mathbf{d}\sigma$ is oriented such that it points away from the interior of R_Σ , and \mathbf{H} and \mathbf{E} are separately arbitrary 3-vectors. We note that if for some 3-scalar V_0 and some 3-vector \mathbf{V} we can write

$$\begin{aligned}\mathbf{H} &= \nabla \times \mathbf{V} - 2V_0 \boldsymbol{\Omega} \\ \mathbf{E} &= \nabla V_0 + V_0 \mathbf{a} - \dot{\mathbf{V}} - \mathbf{S}^* \cdot \mathbf{V}\end{aligned}\quad (31)$$

the complete surface integral on the left of Eq. (30) must vanish, since from Eq. (24) it is given by the line integral with C contracted to a point. Using the commutation relations to differentiate Eq. (31) we can show that for such vectors

$$\nabla \cdot \mathbf{H} = -2\boldsymbol{\Omega} \cdot \mathbf{E} \quad (32)$$

$$(\nabla + \mathbf{a}) \times \mathbf{E} = -\dot{\mathbf{H}} - (Tr\mathbf{S})\mathbf{H} + \mathbf{H} \cdot \mathbf{S}^* \quad (33)$$

so that we have $J^0 \equiv 0$, $\mathbf{J} \equiv 0$, as required for consistency. Four dimensionally, this is just the case $F_{pq} = V_{[p;q]}$, a relation implied, for example, by Maxwell's equations for the electromagnetic field tensor F_{pq} . Indeed, Eqs. (32) and (33) are just the dyadic form of the two Maxwell equations which can be identically satisfied by the introduction of scalar and vector potentials, V_0 and \mathbf{V} , as in Eq. (31). The remaining pair of dyadic Maxwell equations can be obtained by putting $\mathbf{H} \rightarrow \mathbf{E}$ and $\mathbf{E} \rightarrow -\mathbf{H}$ in Eqs. (27) and (28) for J^0 and \mathbf{J} , where $\{J^0, \mathbf{J}\}$ are then identified with the charge and current densities in the local comoving frame. For a local inertial frame ($\mathbf{a}, \boldsymbol{\Omega}, \boldsymbol{\omega}, \mathbf{S}$, all zero), these equations all take their usual form.

When $\mathbf{E} = 0$ Eq. (30) becomes

$$\oint_{\Sigma} \mathbf{H} \cdot \mathbf{d}\sigma = \int_{R_{\Sigma}} \left\{ \nabla \cdot \mathbf{H} - \frac{1}{\phi} \mathbf{A} \cdot [\dot{\mathbf{H}} + (Tr\mathbf{S})\mathbf{H} - \mathbf{H} \cdot \mathbf{S}^*] \right\} dv \quad (34)$$

and again this is actually completely equivalent to the general theorem, as may be verified by putting in this equation, $\mathbf{H} \rightarrow \mathbf{H} + (1/\phi) \mathbf{A} \times \mathbf{E}$, and then eliminating derivatives of ϕ and \mathbf{A} with the potential equations for \mathbf{a} and $\boldsymbol{\Omega}$. Equation (34), then can be adopted as the general form of Gauss' theorem in dyadic notation.

If we consider the integral theorems written for time-like curves and surfaces and for 4-dimensional volumes, we find that the only new information obtained can be summarized in a series of rules for the time differentiation of comoving line, surface, and volume integrals over the quotient space. These relations are

$$\frac{d}{dt} \int_c \mathbf{V} \cdot \mathbf{dr} = \int_c \frac{1}{\phi} (\dot{\mathbf{V}} + \mathbf{S}^* \cdot \mathbf{V}) \cdot \mathbf{dr} \quad (35)$$

$$\frac{d}{dt} \int_{\Sigma} \mathbf{H} \cdot \mathbf{d}\sigma = \int_{\Sigma} \frac{1}{\phi} [\dot{\mathbf{H}} + (Tr\mathbf{S})\mathbf{H} - \mathbf{H} \cdot \mathbf{S}^*] \cdot \mathbf{d}\sigma \quad (36)$$

$$\frac{d}{dt} \int_R \rho dv = \int_R \frac{1}{\phi} [\dot{\rho} + (Tr\mathbf{S})\rho] dv \quad (37)$$

where ρ is an arbitrary 3-scalar and \mathbf{V} and \mathbf{H} are arbitrary 3-vectors.

We can use the second of these equations to derive a relation analogous to Faraday's law in electromagnetism. We put $\mathbf{V} = (1/\phi)\mathbf{E}$ in Eq. (25) and use the potential equation for \mathbf{a} . The result is

$$\oint_c \frac{1}{\phi} \mathbf{E} \cdot \mathbf{dr} = \int_{\Sigma_c} \frac{1}{\phi} \left\{ (\nabla + \mathbf{a}) \times \mathbf{E} - \left[\left(\frac{\mathbf{A} \times \mathbf{E}}{\phi} \right)^{\cdot} + (Tr\mathbf{S}) \left(\frac{\mathbf{A} \times \mathbf{E}}{\phi} \right) - \left(\frac{\mathbf{A} \times \mathbf{E}}{\phi} \right) \cdot \mathbf{S}^* \right] \right\} \cdot \mathbf{d}\sigma \quad (38)$$

Let us now assume that \mathbf{E} satisfies Eq. (33) so that this becomes

$$\oint_c \frac{1}{\phi} \mathbf{E} \cdot \mathbf{dr} = - \int_{\Sigma_c} \frac{1}{\phi} [\dot{\mathbf{F}} + (Tr\mathbf{S})\mathbf{F} - \mathbf{F} \cdot \mathbf{S}^*] \cdot \mathbf{d}\sigma \quad (39)$$

where

$$\mathbf{F} \equiv \mathbf{H} + \frac{1}{\phi} \mathbf{A} \times \mathbf{E}$$

Then using Eq. (36), we have

$$\oint_c \frac{1}{\phi} \mathbf{E} \cdot d\mathbf{r} = - \frac{d}{dt} \int_{\Sigma_c} \left[\mathbf{H} + \frac{1}{\phi} \mathbf{A} \times \mathbf{E} \right] \cdot d\boldsymbol{\sigma} \equiv - \frac{d\Phi}{dt} \quad (40)$$

defining the flux

$$\Phi \equiv \int_{\Sigma_c} \left[\mathbf{H} + \frac{1}{\phi} \mathbf{A} \times \mathbf{E} \right] \cdot d\boldsymbol{\sigma} \quad (41)$$

Likewise we can apply the third of the time derivative rules together with Eq. (34), written for $\mathbf{H} = (1/\phi)\mathbf{J}$, to obtain a conservation law. From Eq. (34) and the potential equations, we have

$$\oint_{\Sigma} \frac{1}{\phi} \mathbf{J} \cdot d\boldsymbol{\sigma} = \int_{R_{\Sigma}} \frac{1}{\phi} \left\{ \nabla \cdot \mathbf{J} + \mathbf{a} \cdot \mathbf{J} - \left[\left(\frac{\mathbf{A} \cdot \mathbf{J}}{\phi} \right)' + (TrS) \left(\frac{\mathbf{A} \cdot \mathbf{J}}{\phi} \right) \right] \right\} dv \quad (42)$$

Now we assume the existence of a 3-scalar ρ such that

$$\nabla \cdot \mathbf{J} + \mathbf{a} \cdot \mathbf{J} + \dot{\rho} + (TrS)\rho = 0 \quad (43)$$

This is the dyadic expression of the equation for a conserved 4-vector, $J^r{}_{;r} = 0$, where the tetrad components are $J^r = \{\rho, \mathbf{J}\}$. Thus,

$$\begin{aligned} \oint_{\Sigma} \frac{1}{\phi} \mathbf{J} \cdot d\boldsymbol{\sigma} &= - \int_{R_{\Sigma}} \frac{1}{\phi} \left[\left(\rho + \frac{1}{\phi} \mathbf{A} \cdot \mathbf{J} \right)' + (TrS) \left(\rho + \frac{1}{\phi} \mathbf{A} \cdot \mathbf{J} \right) \right] dv \\ &= - \frac{d}{dt} \int_{R_{\Sigma}} \left(\rho + \frac{1}{\phi} \mathbf{A} \cdot \mathbf{J} \right) dv \\ &= - \frac{dQ}{dt} \end{aligned} \quad (44)$$

defining the conserved quantity

$$Q \equiv \int_{R_{\Sigma}} \left(\rho + \frac{1}{\phi} \mathbf{A} \cdot \mathbf{J} \right) dv \quad (45)$$

We note also that putting $\rho = 1$ in Eq. (37) gives the time rate-of-change of absolute quotient space volume itself

$$\frac{dQ}{dt} = \int_R \frac{1}{\phi} (TrS) dv \quad (46)$$

and if we adopt proper time on the timelike congruence for the time coordinate, so that

$$\phi = \frac{dt}{d\tau} = 1$$

we get the classical result

$$\frac{dQ}{d\tau} = \int_R (TrS) dv \quad (47)$$

References

1. Estabrook, F. B., and Wahlquist, H. D., "Dyadic Analysis of Spacetime Congruences," *J. Math. Phys.*, Vol. 5, p. 1629, 1964.
2. Wahlquist, H. D., and Estabrook, F. B., *Rigid Motions in Einstein Spaces*, Technical Report 32-868. Jet Propulsion Laboratory, Pasadena, Calif., March 15, 1966.
3. Synge, J. L., *Relativity—The General Theory*, pp. 41–46. North Holland Publishing Co., Amsterdam, 1966.

C. Collisional Quenching of $O(^1D)$, W. B. DeMore

A basic problem in atmospheric chemistry concerns the relative efficiencies of different gases as quenchers of singlet atomic oxygen, $O(^1D)$. In early work, DeMore and Raper (Ref. 1) showed that N_2 is an effective quencher of $O(^1D)$, and also argued on the basis of experimental and semitheoretical considerations (Ref. 2) that O_2 and N_2 should have approximately equal quenching efficiencies. However, this interpretation has been disputed by subsequent investigators (Refs. 3, 4, 5), who have reported that O_2 is a poor quencher relative to N_2 . Most of the latter work has been limited to a technique involving competition between the reactions $O(^1D) + O_3 \rightarrow 2 O_2$ and $O(^1D) + O_2 \rightarrow O(^3P) + O_2$. The reason is that many other methods used for quenching measurements (Refs. 6, 7) are incompatible with the presence of O_2 .

As an aid in resolving this problem, we have developed a new method for measuring relative quenching efficiencies which is suitable for use with O_2 . The technique involves photolysis of O_3 - N_2O -He mixtures in a 40-m White cell with various pressures of added quencher. The $O(^1D)$, produced by O_3 photolysis at 2537 Å, reacts with N_2O to produce some NO, whereas the quenched species, $O(^3P)$, does not. In the presence of O_3 , NO is rapidly converted to N_2O_5 and the N_2O_5 concentration is monitored by its infrared absorption at 1718 cm^{-1} . Thus the amount of N_2O_5 formed is a measure of the steady-state $O(^1D)$ concentration. The basic mixture consisted of 10 torr N_2O , about 0.05 torr O_3 , the added quencher,

and sufficient He to bring the total pressure to 700 torr. It was found that the rate of N_2O_5 formation was independent of the $\text{N}_2\text{O}/\text{He}$ ratio in the range of N_2O pressures employed, indicating negligible $\text{O}(^1\text{D})$ quenching by He.

The N_2O_5 disappeared at a slow and approximately first-order rate, apparently on the walls. Integration of the amount of N_2O_5 formed during a specified radiation interval leads to the following expression for the ratio of peak N_2O_5 concentrations formed without and with added quencher:

$$[\text{N}_2\text{O}_5]_0/[\text{N}_2\text{O}_5] = 1 + k_d[M]/k_{\text{N}_2\text{O}}[\text{N}_2\text{O}] \quad (1)$$

The term on the left-hand side represents the ratio by which N_2O_5 formation was suppressed by the added quencher. The constant $k_{\text{N}_2\text{O}}$ is the total rate constant for $\text{O}(^1\text{D})$ removal by N_2O , and k_d is the quenching rate constant. Thus, a plot of data according to Eq. (1) permits comparison of the rates k_d for different quenchers M .

The results show the expected linearity for different concentrations of a given quencher (Fig. 3). This experimental observation supports the validity of the method for comparing efficiencies of different quenchers. Note also that a mixture of O_2 and N_2 in equal amounts gave the same quenching as that predicted for the sum of the individual components, thus arguing against the possibility of any unique chemical effects of O_2 on the method. From the slopes, the following relative quenching efficiencies are found: $\text{CO}_2:\text{CO}:\text{N}_2:\text{O}_2 = 1.00:0.41:0.23:0.23$. Quenching by Ar was almost negligible. Experiments with all the He replaced by Ar gave an upper limit of 5×10^{-4} for Ar relative to CO_2 .

These experiments confirm our earlier conclusion that O_2 is an effective quencher of $\text{O}(^1\text{D})$. The reason for the conflicting results arising from other workers (Refs. 3, 4, 5) probably lies in the fact that the reaction $\text{O}(^1\text{D}) + \text{O}_3 \rightarrow 2 \text{O}_2$ is not a reliable index to the steady-state $\text{O}(^1\text{D})$ concentration except under carefully controlled conditions. Competing reactions which decompose O_3 can easily be mistaken for the above $\text{O}(^1\text{D})$ reaction, thus

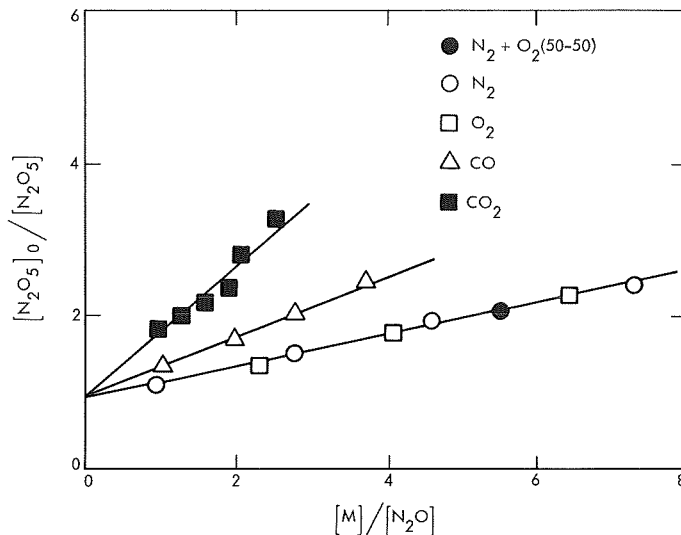


Fig. 3. Suppression of N_2O_5 formation resulting from quenching of $\text{O}(^1\text{D})$ by added gases (quenching efficiencies are proportional to slopes)

leading to an erroneous conclusion that the $\text{O}(^1\text{D})$ has not been quenched.

The present work also confirms that both CO_2 and CO are effective quenchers of $\text{O}(^1\text{D})$. These results are relevant to the problem of CO_2 photochemistry in the atmospheres of Mars and Venus.

References

1. DeMore, W., and Raper, O. F., *J. Chem. Phys.*, Vol. 37, p. 2048, 1962.
2. DeMore, W., and Raper, O. F., *Astrophys. J.*, Vol. 139, p. 1381, 1964.
3. Warneck, P., and Sullivan, J. O., *Ber. Bunsen. physik. Chemie*, 1968, p. 159.
4. Snelling, D. R., and Bair, E. J., *J. Chem. Phys.*, Vol. 48, p. 5737, 1968.
5. Clerc, M., Reiffsteck, A., and Lesigne, B., *J. Chem. Phys.*, Vol. 50, p. 3721, 1969.
6. Yamazaki, H., and Cvetanovic, R. J., *J. Chem. Phys.*, Vol. 41, p. 3703, 1964.
7. Preston, K. F., and Cvetanovic, R. J., *J. Chem. Phys.*, Vol. 45, p. 2888, 1966.

VI. Communications Elements Research

TELECOMMUNICATIONS DIVISION

A. Spacecraft Antenna Research: Fields in a Cylindrically Inhomogeneous Plasma Excited by an Arbitrarily Oriented Dipole Antenna, R. Woo

1. Introduction

In SPS 37-56, Vol. III, pp. 92-97, the use of geometrical optics to study radiation from a dipole in an inhomogeneous plasma was formulated. By using the reciprocity theorem the problem was reduced to the simpler one of solving the fields for a plane wave incident on the plasma. This method yields only the far-field radiation patterns. In this article, the reciprocity theorem will not be used. Instead, the problem is reformulated and then solved completely. The geometrical optics approximation is still used to represent the fields. As a result, the fields excited by the dipole everywhere are readily obtained.

2. Geometrical Optics Formulation

Let the electric field be written in the following form:

$$\mathbf{E}(\rho, \phi, z) = A(\rho, \phi, z) \hat{\mathbf{P}}(\rho, \phi, z) \exp[-jk\psi(\rho, \phi, z)] \quad (1)$$

where A is the amplitude, $\hat{\mathbf{P}}$ is the unit polarization vector and ψ is the phase. According to the geometrical optics

approximation, if the spatial change per wavelength of refractive index, absorption loss, and field amplitude are small, then ψ satisfies the eikonal equation while A and $\hat{\mathbf{P}}$ satisfy the vector transport equation (Refs. 1-3).¹

a. Phase. The phase function ψ satisfies the eikonal equation

$$\nabla\psi \cdot \nabla\psi = n^2 \quad (2)$$

where n is the index of refraction. For an isotropic plasma, n is given by

$$n^2 = 1 - \left(\frac{\omega_p}{\omega}\right)^2 \quad (3)$$

where ω_p is the electron plasma frequency. For the plasma configuration considered in this article, n is a function of radius only. Equation (2) can therefore be written

$$\left(\frac{\partial\psi}{\partial\rho}\right)^2 + \frac{1}{\rho^2}\left(\frac{\partial\psi}{\partial\phi}\right)^2 + \left(\frac{\partial\psi}{\partial z}\right)^2 = n^2(\rho) \quad (4)$$

¹This solution corresponds to the first term of the asymptotic series solution. For discussion, see SPS 37-56, Vol. III, pp. 92-97 and Refs. 1-3.

The complete integral of Eq. (4) is obtained by using the method of separation of variables (Ref. 4). Assume ψ can be written in the form

$$\psi(\rho, \phi, z) = f_1(\rho) + f_2(\phi) + f_3(z) \quad (5)$$

then

$$\frac{df_3(z)}{dz} = \text{constant} = a$$

$$\frac{df_2(\phi)}{d\phi} = \text{constant} = b$$

$$\frac{df_1(\rho)}{d\rho} = \left[n^2(\rho) - a^2 - \frac{b^2}{\rho^2} \right]^{1/2}$$

Thus,

$$\begin{aligned} \psi(\rho, \phi, z) = & a(z - z_0) + b(\phi - \phi_0) \\ & + \int_{\rho_0}^{\rho} \left[n^2(\rho) - a^2 - \frac{b^2}{\rho^2} \right]^{1/2} d\rho \end{aligned} \quad (6)$$

where the phase at (ρ_0, ϕ_0, z_0) is zero. When combined with the ray Eqs. (7) and (8), the complete integral Eq. (6), in fact, yields the general solution to the eikonal equation (Ref. 4).

The equations for the rays can be obtained from the wave fronts by using the theorem of Jacobi (Refs. 1 and 2). According to this theorem, the ray equations are given by

$$\frac{\partial \psi}{\partial a} = \text{constant} = \alpha$$

$$\frac{\partial \psi}{\partial b} = \text{constant} = \beta$$

Letting $\alpha = z_0$ and $\beta = \phi_0$, the ray equations are

$$z = z_0 + \int_{\rho_0}^{\rho} \frac{a}{\left[n^2(\rho) - a^2 - \frac{b^2}{\rho^2} \right]^{1/2}} d\rho \quad (7)$$

$$\phi = \phi_0 + \int_{\rho_0}^{\rho} \frac{b/\rho^2}{\left[n^2(\rho) - a^2 - \frac{b^2}{\rho^2} \right]^{1/2}} d\rho \quad (8)$$

These two equations determine the ray for given values of a and b . The constants a and b can be interpreted in terms of the ray geometry by considering Eqs. (7) and (8)

along with the fact that

$$ds^2 = d\rho^2 + \rho^2 d\phi^2 + dz^2 \quad (9)$$

where ds is the differential path length along the ray. From these equations it can be shown that

$$a = n(\rho) \frac{dz}{ds} \quad (10)$$

$$b = n(\rho) \rho^2 \frac{d\phi}{ds} \quad (11)$$

If the angle the ray makes with the z axis is denoted γ while the angle the projection of the ray on the x - y plane makes with the ρ axis is denoted ξ (Fig. 1), then

$$a = n(\rho) \cos \gamma \quad (12)$$

$$b = n(\rho) \rho \sin \gamma \sin \xi \quad (13)$$

Note from Fig. 1 that if ϕ' is the angle the projection of the ray on the x - y plane makes with the x axis, then $\phi' = \xi + \phi$. The quantities represented by Eqs. (12) and (13) are preserved throughout the path of the ray and constitute the generalization of Snell's law.

b. Amplitude. The amplitude A along a ray is described by the transport equation. The solution for A in terms of general geometrical parameters is well known (Refs. 1, 2, 3, and 5). Let the subscripts 1 and 2 denote evaluation at

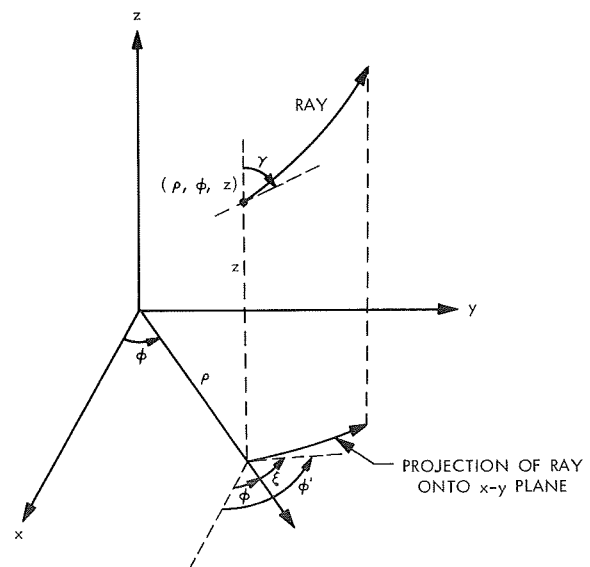


Fig. 1. Definitions of γ and ξ

points (ρ_1, ϕ_1, z_1) and (ρ_2, ϕ_2, z_2) , respectively. Then²

$$A_2 = A_1 \left(\frac{n_1}{n_2} \frac{d\sigma_1}{d\sigma_2} \right)^{1/2} \quad (14)$$

where $d\sigma$ is the cross-sectional area of a tube of rays. Equation (14) is merely the statement that energy is conserved in a tube of rays. If the medium is lossy, Eq. (14) must be multiplied by the factor

$$\exp \left[- \int_{s_1}^{s_2} \frac{\sigma'}{2n} \left(\frac{\mu_0}{\epsilon_0} \right)^{1/2} ds \right] \quad (15)$$

where σ' is the conductivity of the medium. According to Eq. (14), amplitude is given once the cross-sectional area of a tube of rays is found. Figure 2 illustrates how this latter quantity is computed for the configuration at hand. Let dA , the small element of area on the $\rho = \text{constant}$ surface in Fig. 2, define the tube of rays. Then, the cross-sectional area of this tube of rays is

$$d\sigma = dA \hat{\rho} \cdot \hat{T} \quad (16)$$

where $\hat{\rho}$ is the unit vector in the ρ direction and \hat{T} is the unit vector tangent to the ray and oriented in the direc-

²In SPS 37-56, Vol. III, p. 95,

$$A_2 = A_1 \exp \left(- \int_{s_1}^{s_2} \frac{\nabla^2 \psi}{2n} ds \right)$$

That this expression is equivalent to Eq. (14) can be verified by application of Gauss's divergence theorem.

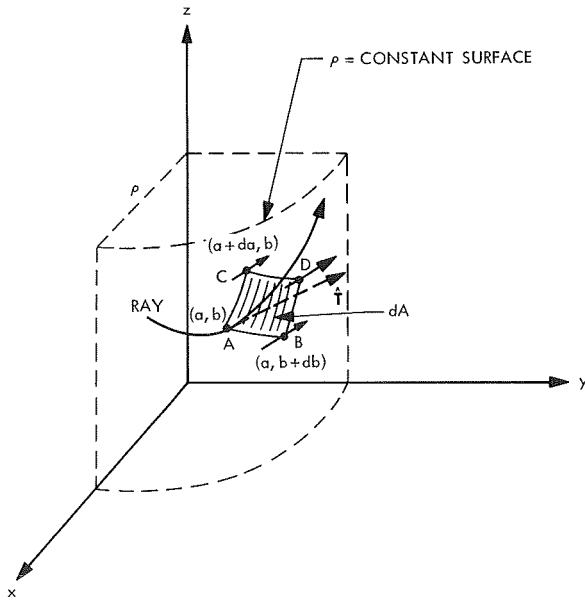


Fig. 2. Calculation of $d\sigma$

tion of the ray. Since the rays are orthogonal trajectories to the wave fronts for an isotropic medium, Eq. (2) yields

$$\hat{T} = \frac{\nabla \psi}{n} \quad (17)$$

To determine dA , note in Fig. 2 that vectors \vec{AB} and \vec{AC} are

$$\vec{AB} = \rho \frac{\partial \phi}{\partial b} db \hat{\phi} + \frac{\partial z}{\partial b} db \hat{z} \quad (18)$$

$$\vec{AC} = \rho \frac{\partial \phi}{\partial a} da \hat{\phi} + \frac{\partial z}{\partial a} da \hat{z} \quad (19)$$

where $\hat{\phi}$ and \hat{z} are the unit vectors in the ϕ and z directions, respectively. From Eqs. (7) and (8),

$$\frac{\partial z}{\partial a} = - \frac{\partial^2}{\partial a^2} \int_{\rho_0}^{\rho} \left[n^2(\rho) - a^2 - \frac{b^2}{\rho^2} \right]^{1/2} d\rho = - \psi_{aa} \quad (20)$$

$$\frac{\partial \phi}{\partial b} = - \frac{\partial^2}{\partial b^2} \int_{\rho_0}^{\rho} \left[n^2(\rho) - a^2 - \frac{b^2}{\rho^2} \right]^{1/2} d\rho = - \psi_{bb} \quad (21)$$

$$\frac{\partial \phi}{\partial a} = \frac{\partial z}{\partial b} = - \frac{\partial^2}{\partial a \partial b} \int_{\rho_0}^{\rho} \left[n^2(\rho) - a^2 - \frac{b^2}{\rho^2} \right]^{1/2} d\rho = - \psi_{ab} \quad (22)$$

Now,

$$dA = |\vec{AC} \times \vec{AB}|$$

Thus,

$$dA = \rho (\psi_{aa}\psi_{bb} - \psi_{ab}^2) da db \quad (23)$$

From Eqs. (6) and (17),

$$\hat{T} = \frac{1}{n} \left[\left(n^2 - a^2 - \frac{b^2}{\rho^2} \right)^{1/2} \hat{\rho} + \frac{b}{\rho} \hat{\phi} + a \hat{z} \right] \quad (24)$$

so that

$$d\sigma = \frac{1}{n} \left[n^2 - a^2 - \frac{b^2}{\rho^2} \right]^{1/2} \rho (\psi_{aa}\psi_{bb} - \psi_{ab}^2) da db \quad (25)$$

If $\psi_{aa}\psi_{bb} - \psi_{ab}^2$ is denoted W , then

$$\frac{A_2}{A_1} = \frac{\rho_1^{1/2}}{\rho_2^{1/2}} \frac{W_1^{1/2}}{W_2^{1/2}} \left[\frac{n^2(\rho_1) - a^2 - \frac{b^2}{\rho_1^2}}{n^2(\rho_2) - a^2 - \frac{b^2}{\rho_2^2}} \right]^{1/4} \quad (26)$$

c. Polarization. For polarization, solution of the transport equations gives the following results for the unit polarization vector $\hat{\mathbf{P}}$. If $\hat{\mathbf{P}}_1$ is expressed as

$$\hat{\mathbf{P}}_1 = \hat{\mathbf{N}}_1 \cos \delta_1 + \hat{\mathbf{S}}_1 \sin \delta_1 \quad (27)$$

then $\hat{\mathbf{P}}_2$ is given by

$$\hat{\mathbf{P}}_2 = \hat{\mathbf{N}}_2 \cos \delta_2 + \hat{\mathbf{S}}_2 \sin \delta_2 \quad (28)$$

where

$$\delta_2 = \delta_1 - \int_{s_1}^{s_2} \frac{1}{\tau} ds \quad (29)$$

$\hat{\mathbf{N}}$ and $\hat{\mathbf{S}}$ are the respective normal and binormal vectors of the ray while τ is the torsion of the ray. $\hat{\mathbf{N}}$ and $\hat{\mathbf{S}}$ are defined in terms of the tangent $\hat{\mathbf{T}}$ as follows (Ref. 5):

$$\hat{\mathbf{N}} = \rho_c \frac{d\hat{\mathbf{T}}}{ds} \quad (30)$$

$$\hat{\mathbf{S}} = \hat{\mathbf{T}} \times \hat{\mathbf{N}} \quad (31)$$

where the radius of curvature ρ_c is given by $1/\rho_c = |d\hat{\mathbf{T}}/ds|$. But,

$$\frac{d\hat{\mathbf{T}}}{ds} = (\hat{\mathbf{T}} \cdot \nabla) \hat{\mathbf{T}} \quad (32)$$

Thus, the combination of Eqs. (24), (30), (31), and (32) yields

$$\begin{aligned} \hat{\mathbf{N}} = & \frac{1}{n} \left(a^2 + \frac{b^2}{\rho^2} \right)^{1/2} \hat{\boldsymbol{\rho}} - \frac{b}{n\rho} \left(\frac{n^2}{a^2 + \frac{b^2}{\rho^2}} - 1 \right)^{1/2} \hat{\boldsymbol{\phi}} \\ & - \frac{a}{n} \left(\frac{n^2}{a^2 + \frac{b^2}{\rho^2}} - 1 \right)^{1/2} \hat{\mathbf{z}} \end{aligned} \quad (33)$$

$$\hat{\mathbf{S}} = \frac{a}{\left(a^2 + \frac{b^2}{\rho^2} \right)^{1/2}} \hat{\boldsymbol{\phi}} - \frac{b/\rho}{\left(a^2 + \frac{b^2}{\rho^2} \right)^{1/2}} \hat{\mathbf{z}} \quad (34)$$

Torsion τ is related to $\hat{\mathbf{N}}$ and $\hat{\mathbf{S}}$ through the Frenet-Serret formulas (Ref. 6).

$$\frac{1}{\tau} = \frac{d\hat{\mathbf{N}}}{ds} \cdot \hat{\mathbf{S}} \quad (35)$$

But

$$\frac{d\hat{\mathbf{N}}}{ds} = (\hat{\mathbf{T}} \cdot \nabla) \hat{\mathbf{N}} \quad (36)$$

Therefore, the combination of Eqs. (24) and (33)–(36) yields

$$\frac{1}{\tau} = \frac{ab}{n^2\rho} \frac{n^2 - a^2 - \frac{b^2}{\rho^2}}{a^2 + \frac{b^2}{\rho^2}} \quad (37)$$

From Eqs. (9), (10), and (11)

$$\frac{ds}{d\rho} = \frac{n}{\left(n^2 - a^2 - \frac{b^2}{\rho^2} \right)^{1/2}} \quad (38)$$

and Eq. (29) becomes

$$\delta_2 = \delta_1 - \int_{\rho_1}^{\rho_2} \frac{ab}{n\rho^2} \frac{\left(n^2 - a^2 - \frac{b^2}{\rho^2} \right)^{1/2}}{a^2 + \frac{b^2}{\rho^2}} d\rho \quad (39)$$

The geometrical optics representation is now complete. Phase is given by Eq. (6), amplitude by Eq. (26), and polarization by Eqs. (27), (28), (33), (34), and (39). Thus, if the fields are known at (ρ_1, ϕ_1, z_1) , the respective fields at (ρ_2, ϕ_2, z_2) may be calculated. All that remains is to include the source.

d. Inclusion of source. To include the source, the dipole antenna is assumed to be located at (ρ_0, ϕ_0, z_0) in a homogeneous medium whose index of refraction is the same as that where the antenna is located in the inhomogeneous medium. Fields are obtained at (ρ_1, ϕ_1, z_1) , which is assumed to be located physically close to (ρ_0, ϕ_0, z_0) yet far enough to be in the far field. For a dipole antenna oriented in the direction $\hat{\mathbf{i}}$ and located in a medium whose

index of refraction is n_0 , \mathbf{E} can be written in terms of the electric Hertz potential π

$$\mathbf{E} = \nabla \times \nabla \times \pi \quad (40)$$

$$\pi = \frac{Il}{j\omega\epsilon_0} \frac{\exp(-jkn_0 r)}{4\pi r} \hat{\mathbf{i}} \quad (41)$$

where r is the distance from the antenna and Il is the dipole moment. In the far field, $\nabla \rightarrow -jkn_0 \mathbf{r}$. Therefore,

$$\begin{aligned} \mathbf{E} &= j\omega\mu Il \frac{\exp(-jkn_0 r)}{4\pi r} [\hat{\mathbf{r}} \times (\hat{\mathbf{r}} \times \hat{\mathbf{i}})] \\ &= j\omega\mu Il \frac{\exp(-jkn_0 r)}{4\pi r} [(\hat{\mathbf{r}} \cdot \hat{\mathbf{i}}) \hat{\mathbf{r}} - \hat{\mathbf{i}}] \end{aligned} \quad (42)$$

Let $\mathbf{Q} = (\hat{\mathbf{r}} \cdot \hat{\mathbf{i}}) \hat{\mathbf{r}} - \hat{\mathbf{i}}$. Then $\hat{\mathbf{P}} = \mathbf{Q}/|\mathbf{Q}|$ and

$$\mathbf{E}_1 = j\omega\mu Il \frac{\exp(-jkn_0 r_1)}{4\pi r_1} |\mathbf{Q}_1| \hat{\mathbf{P}}_1 \quad (43)$$

Comparing Eq. (43) with Eq. (1),

$$A_1 = j\omega\mu Il \frac{|\mathbf{Q}_1|}{4\pi r_1} \quad (44)$$

Note too that the phase term in Eq. (43) agrees with Eq. (6). Now, W_1 in Eq. (26) can be evaluated keeping in mind the comments made above and after eliminating ρ_1

$$\mathbf{E}_2 = j\omega\mu Il \frac{|\mathbf{Q}_1|}{4\pi} \frac{1}{\rho_0^{1/2} \left(n_0^2 - a^2 - \frac{b^2}{\rho_0^2} \right)^{1/4}} \frac{1}{\rho_2^{1/2} W_2^{1/2}} \frac{1}{\left(n_2^2 - a^2 - \frac{b^2}{\rho_2^2} \right)^{1/4}} \exp(-jk\psi) \hat{\mathbf{P}}_2 \quad (45)$$

where ψ is given in Eq. (6). It must be remembered that if the plasma is lossy, Eq. (45) must be multiplied by the factor in Eq. (15). To determine $\hat{\mathbf{P}}_2$ it is convenient to use cartesian coordinates. It is assumed that

$$\hat{\mathbf{i}} = A\hat{\mathbf{x}} + B\hat{\mathbf{y}} + C\hat{\mathbf{z}} \quad (46)$$

where A , B , and C are the direction cosines. Making the transformation

$$\left. \begin{aligned} \hat{\mathbf{p}} &= \cos \phi \hat{\mathbf{x}} + \sin \phi \hat{\mathbf{y}} \\ \hat{\mathbf{q}} &= -\sin \phi \hat{\mathbf{x}} + \cos \phi \hat{\mathbf{y}} \\ \hat{\mathbf{z}} &= \hat{\mathbf{z}} \end{aligned} \right\} \quad (47)$$

in Eq. (24), $\hat{\mathbf{r}}$ can be written

$$\hat{\mathbf{r}} = (\alpha \cos \phi - \beta \sin \phi) \hat{\mathbf{x}} + (\alpha \sin \phi + \beta \cos \phi) \hat{\mathbf{y}} + \eta \hat{\mathbf{z}} \quad (48)$$

where

$$\begin{aligned} \alpha &= \frac{1}{n_0} \left(n_0^2 - a^2 - \frac{b^2}{\rho_1^2} \right)^{1/2} \\ \beta &= \frac{1}{n_0} \frac{b}{\rho_1} \\ \eta &= \frac{a}{n_0} \end{aligned}$$

Using Eqs. (46) and (48), $\hat{\mathbf{P}}_1$ can now be readily determined. $\hat{\mathbf{N}}_1$ and $\hat{\mathbf{S}}_1$ are given in Eqs. (33) and (34), and these may also be written in terms of cartesian coordinates by making the transformation in Eq. (47). With $\hat{\mathbf{P}}_1$, $\hat{\mathbf{N}}_1$ and $\hat{\mathbf{S}}_1$ determined, δ_1 can be obtained from Eq. (27) by either

$$\cos \delta_1 = \hat{\mathbf{P}}_1 \cdot \hat{\mathbf{N}}_1 \quad (49)$$

or

$$\sin \delta_1 = \hat{\mathbf{P}}_1 \cdot \hat{\mathbf{S}}_1 \quad (50)$$

Finally, $\hat{\mathbf{P}}_2$ is determined from Eqs. (28) and (39). The information necessary to compute the electric field along a ray is now complete.

To determine the fields excited by the dipole, a group of rays emanating from the dipole are traced. At any point along each ray, \mathbf{E} is given in Eq. (45). The values of a and b to be used with each ray are determined from the initial conditions of the ray. If one merely wants to calculate the field at one given point, the process is more difficult. The difficulty arises when finding the ray or rays that connect the given point with the dipole. Analytically, this corresponds to finding the values of a and b associated with the ray or rays connecting the two points and involves the solution of the integral Eqs. (7) and (8). Still, such a computation could be carried out on a computer.

It should be pointed out that the scheme described above is not limited to dipole antennas. If the patterns of any antenna located in a homogeneous medium $n = n_0$ are known, the fields at (ρ_1, ϕ_1, z_1) and, consequently, (ρ_2, ϕ_2, z_2) can be determined.

3. Radiation Patterns

In this subsection, the procedure for calculating the radiation fields is outlined. The antenna is assumed to be located at (ρ_0, ϕ_0, z_0) in the cylindrical plasma whose electron density decays to zero far from the axis. The radiation fields on the surface of a sphere whose radius is R and whose center corresponds to the origin will be calculated. R is chosen to be large so that the sphere is in the far field. Points on the sphere are denoted by subscript 2. Rays emanating at (ρ_0, ϕ_0, z_0) and making the angles $\gamma = \gamma_0$ and $\xi = \xi_0$ at (ρ_0, ϕ_0, z_0) are considered. Because of symmetry, γ_0 ranges from 0 to $\pi/2$ while ξ_0 from 0 to π . For each ray, the following series of calculations is made to determine the direction of radiation that in spherical coordinates is denoted $(\theta_{\text{rad}}, \phi_{\text{rad}})$. The respective values of a and b are determined according to Eqs. (11) and (12). Since the radiated wave is in free space, n_2 is unity and

$$\theta_{\text{rad}} = \gamma_0 = \cos^{-1} a \quad (51)$$

ρ_2 is then given by $R \sin \theta_{\text{rad}}$. Next, ϕ_2 and z_2 are computed using Eqs. (7) and (8). From Eq. (13), ξ_2 may also be obtained and ϕ_{rad} is given by

$$\phi_{\text{rad}} = \phi_2 + \xi_2 \quad (52)$$

With the direction of radiation $(\theta_{\text{rad}}, \phi_{\text{rad}})$ obtained, all that remains is to compute \mathbf{E}_2 according to Eq. (45). This may be done by substituting the computed values of ρ_2 , ϕ_2 , and z_2 into the respective expressions. It should be pointed out that although the procedure is straightforward, the integrals Eq. (7), Eq. (8), W_2 , and Eq. (39) may

have to be numerically evaluated on the computer. Also, turning points may occur in the plasma. The condition for existence of turning points is that $[a^2 + (b^2/\rho^2)]$ be greater than the minimum of $n(\rho)$. These turning points give rise to refracted fields that may be calculated according to the method used by B. D. Seckler and J. B. Keller (Ref. 5).

When the antenna is located on the z axis, the symmetry allows explicit expressions for the radiation patterns if the above procedure is followed. The results for three cases are given below. The plasma is assumed to be lossless. In all cases, $\cos^{-1} n_{\text{max}} \leq \theta_{\text{rad}} \leq (\pi/2)$ and $0 \leq \phi_{\text{rad}} \leq 2\pi$ where n_{max} is the index of refraction corresponding to the peak electron density in the plasma. For $0 \leq \theta_{\text{rad}} \leq \cos^{-1} n_{\text{max}}$, the fields are zero. n_{ant} is the index of refraction at the location of the antenna.

a. Vertical dipole. Dipole oriented in the z direction
 $\hat{\mathbf{i}} = \hat{\mathbf{z}}$

$$\left. \begin{aligned} E_\theta &= j\omega\mu Il \left(1 - \frac{\cos^2 \theta_{\text{rad}}}{n_{\text{ant}}}\right)^{1/2} \frac{\exp(-jkR)}{R} \\ E_\phi &= 0 \end{aligned} \right\} \quad (53)$$

b. Horizontal dipole. Dipole oriented in the x direction
 $\hat{\mathbf{i}} = \hat{\mathbf{x}}$

$$\left. \begin{aligned} E_\theta &= -j\omega\mu Il \frac{\cos \theta_{\text{rad}}}{n_{\text{ant}}} \cos \phi_{\text{rad}} \frac{\exp(-jkR)}{4\pi R} \\ E_\phi &= j\omega\mu Il \sin \phi_{\text{rad}} \frac{\exp(-jkR)}{4\pi R} \end{aligned} \right\} \quad (54)$$

c. Turnstile antenna. The antenna consists of dipoles oriented in the x and y directions and located $\lambda/4$ above an infinite ground plane. The dipole in the y direction is fed -90 deg out of phase with respect to the dipole in the x direction. The antenna is therefore right-hand circularly polarized.

$$\left. \begin{aligned} E_{\text{RHCP}} &= \omega\mu Il 2^{1/2} \left(1 + \frac{\cos \theta_{\text{rad}}}{n_{\text{ant}}}\right) \sin\left(\frac{\pi}{2} \cos \theta_{\text{rad}}\right) \exp(-j\phi_{\text{rad}}) \frac{\exp(-jkR)}{4\pi R} \\ E_{\text{LHCP}} &= \omega\mu Il 2^{1/2} \left(1 - \frac{\cos \theta_{\text{rad}}}{n_{\text{ant}}}\right) \sin\left(\frac{\pi}{2} \cos \theta_{\text{rad}}\right) \exp(-j\phi_{\text{rad}}) \frac{\exp(-jkR)}{4\pi R} \end{aligned} \right\} \quad (55)$$

where RHCP and LHCP are right- and left-hand circular polarization, respectively. It is interesting to note that Eqs. (53), (54), and (55) are independent of the plasma electron density profile. This is a consequence of choosing the observation point far away (large R).

Shown in Figs. 3-6 are patterns obtained using Eq. (55) superimposed on the harmonic series solutions for a cylindrically stratified plasma obtained in a previous study (Ref. 7). In each figure the upper patterns are RHCP while the lower patterns are LHCP. Both column and shell con-

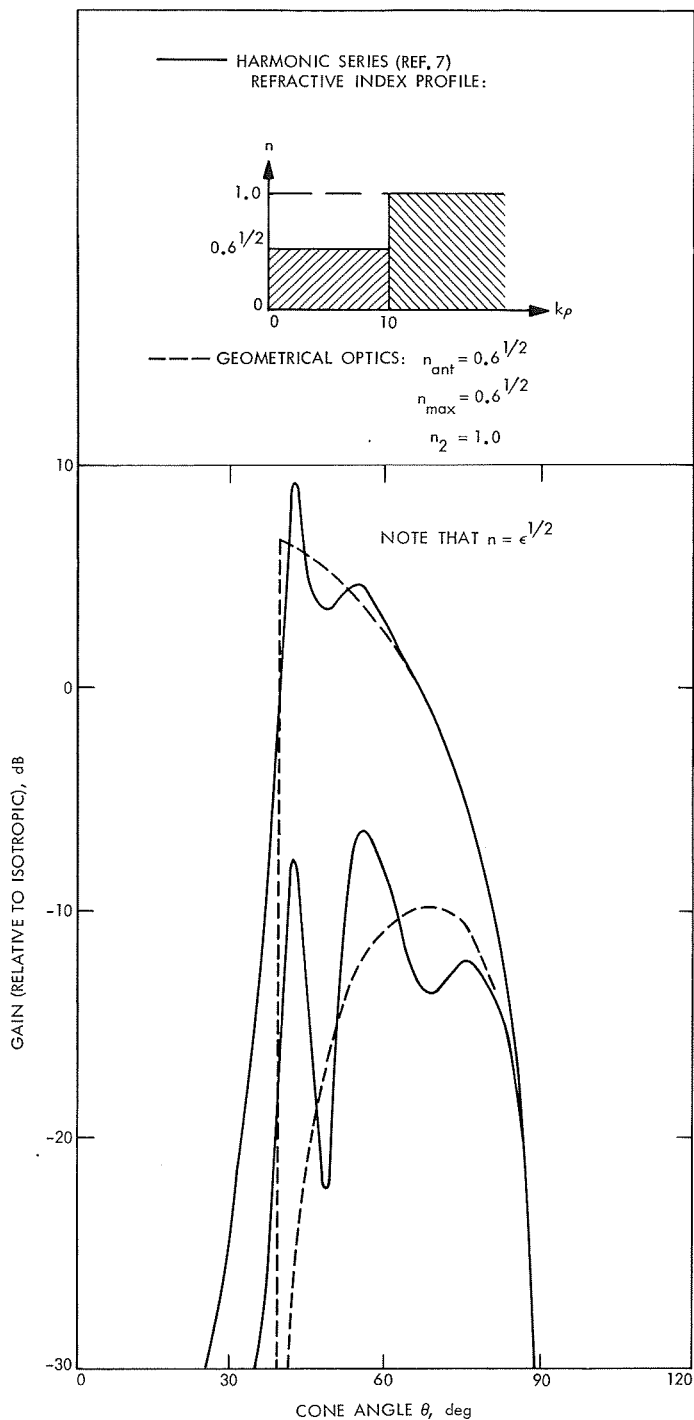


Fig. 3. Patterns for column configuration
($ka = 10$)

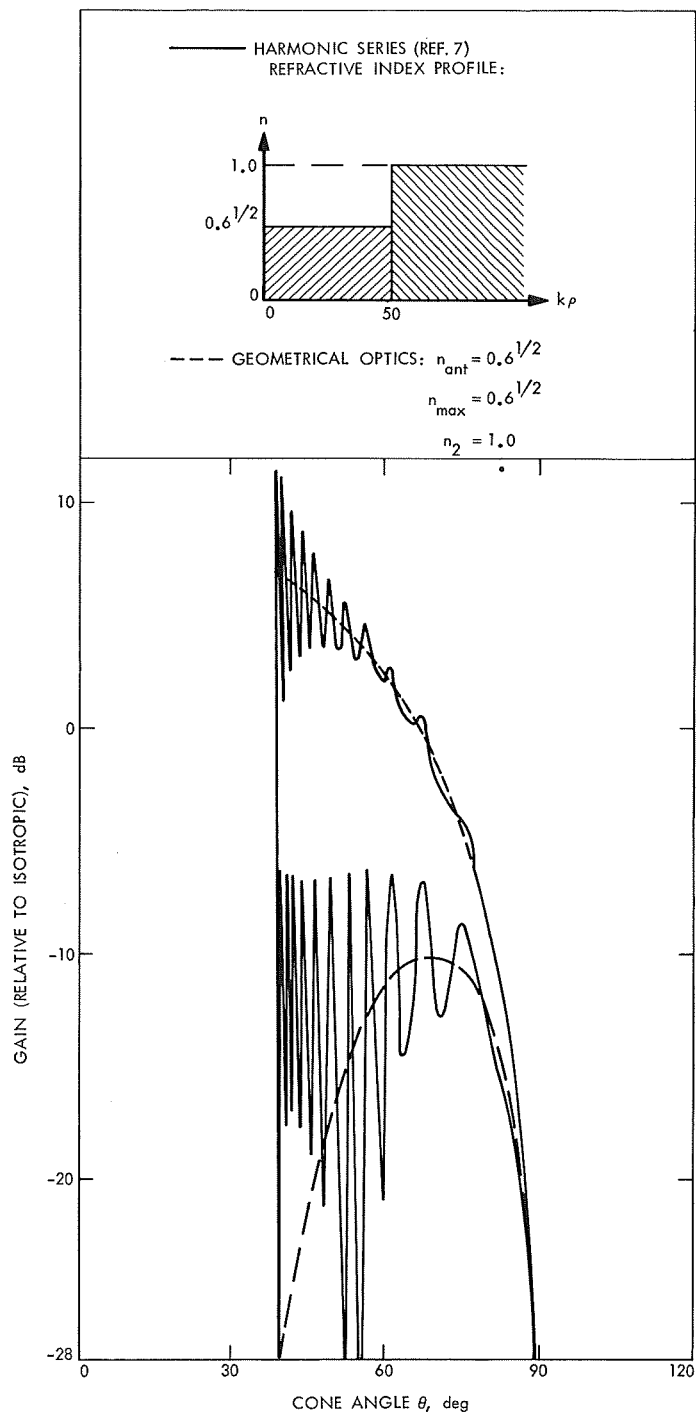


Fig. 4. Patterns for column configuration
($ka = 50$)

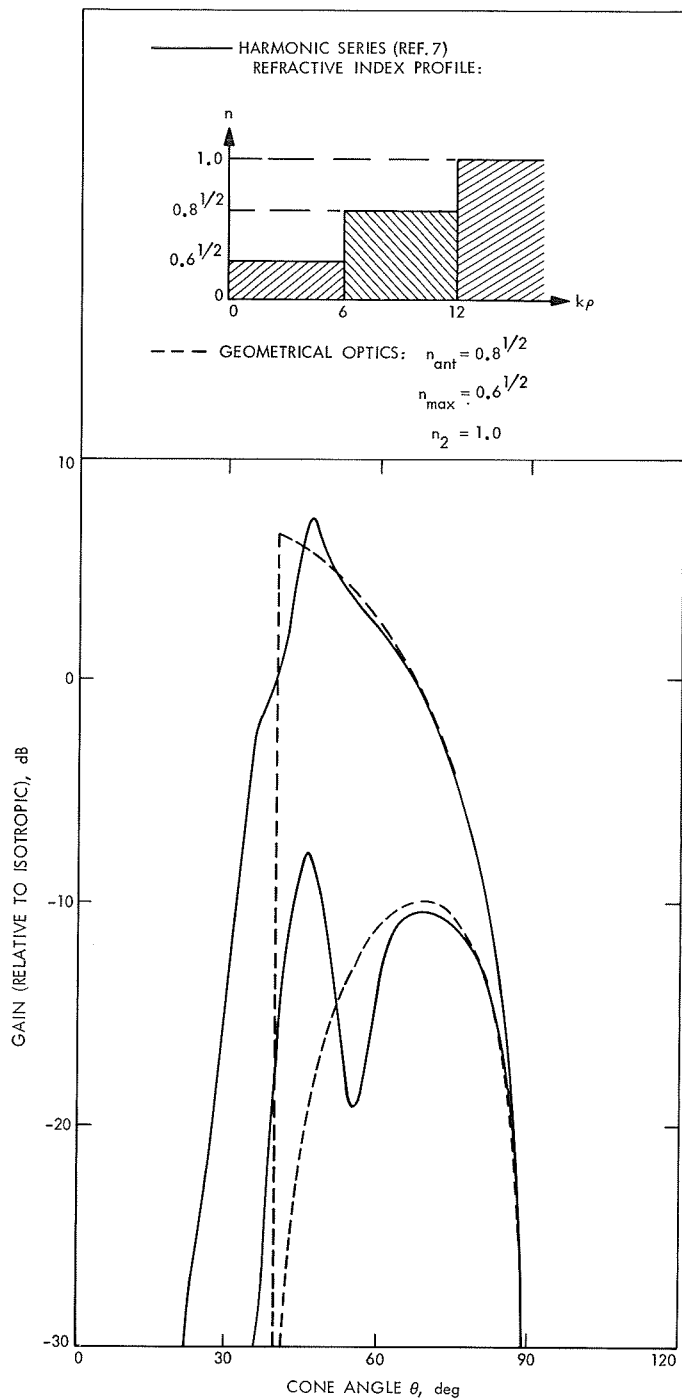


Fig. 5. Patterns for shell configuration
(far-wake approximation)

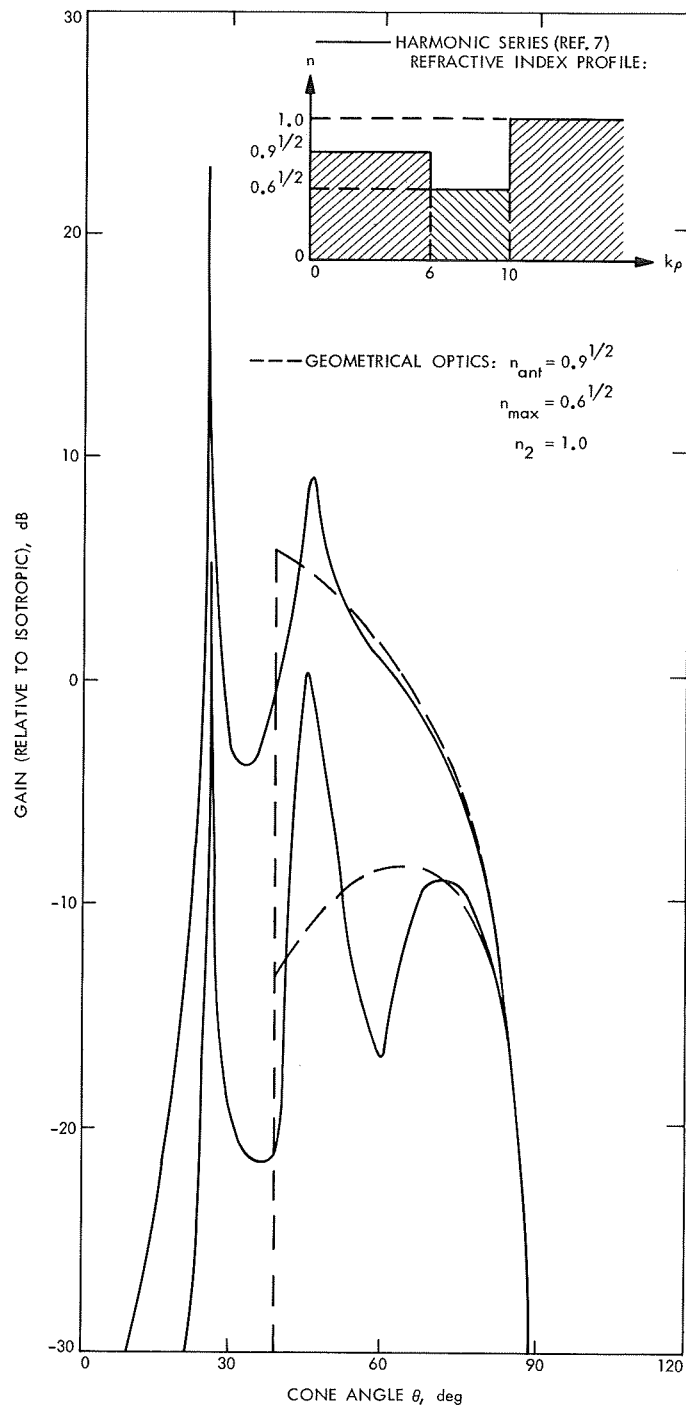


Fig. 6. Patterns for shell configuration
(near-wake approximation)

figurations are included and, in the case of the shell configuration, both near and far wake cases are shown. The refractive index (electron density) profiles are shown in Figs. 3–6. It must be emphasized again that the geometrical optics patterns are independent of electron density profile and depend only on n_{ant} and n_{max} . The geometrical optics patterns are seen to correspond to the “average” of the harmonic series solutions. This is not surprising since it was seen in Ref. 7 that the harmonic series solutions became smoother as the number of plasma layers was increased. The correspondence is remarkable in view of the fact that the expressions in Eq. (55) are much simpler than the harmonic series solutions. Note that the leaky-wave beam in Fig. 6 is not predicted by the geometrical optics representation.

It should be pointed out that an alternate method, applicable to any on-axis antenna may be used to derive Eqs. (53), (54), and (55). This method can be used whenever explicit expressions for the radiation patterns in a homogeneous medium with $n = n_{\text{ant}}$ are available. Let these patterns be written in functional form

$$f(\theta, \phi) \quad (56)$$

If the antenna is immersed in an inhomogeneous plasma whose index of refraction is slowly varying and decays to zero far from the axis, its patterns can be obtained by applying Snell's law. According to Eqs. (12) and (13),

$$\left. \begin{aligned} a &= n(\rho) \cos \theta \\ b &= 0 \end{aligned} \right\} \quad (57)$$

Substituting Eq. (57) into Eq. (56), the resulting patterns are

$$f \left[\cos^{-1} \left(\frac{\cos \theta_{\text{rad}}}{n_{\text{ant}}} \right)_1 \phi_{\text{rad}} \right] \quad (58)$$

where

$$\begin{aligned} 0 &\leq \phi_{\text{rad}} \leq 2\pi \\ 0 &\leq \theta_{\text{rad}} \leq \cos^{-1} n_{\text{max}} \end{aligned}$$

4. Conclusion

The fields excited by a dipole in an inhomogeneous plasma have been solved using the geometrical optics approximation. Radiation patterns for a turnstile antenna have been compared with the harmonic series solutions obtained in a previous study. The correspondence was

seen to be good. This method can be applied to other plasma geometries as well. The condition for application is that the eikonal equation can be solved.

References

1. Luneburg, R. K., *Mathematical Theory of Optics*, University of California Press, Berkeley, Calif., 1964.
2. Kline, M., and Kay, I. W., *Electromagnetic Theory and Geometrical Optics*, Interscience Publishers, New York, 1965.
3. Born, M., and Wolf, E., *Principles of Optics*, pp. 109–132, Pergamon Press, Oxford, London, 1965.
4. Courant, R., and Hilbert, D., *Methods of Mathematical Physics, Vol. II*, Interscience Publishers, New York, 1962.
5. Seckler, B. D., and Keller, J. B., “Geometrical Theory of Diffraction in Inhomogeneous Media,” *J. Acoust. Soc.*, Vol. 31, pp. 192–205, Feb. 1959.
6. Struik, D. J., *Differential Geometry*, Addison-Wesley, Reading, Mass., 1957.
7. Woo, R., and Ishimaru, A., “Radiation from a Circularly Polarized Antenna through the Ionized Wake of a Mars-Entry Capsule,” *IEEE Trans. Antennas and Propagation*, AP-17, pp. 488–495, July 1969.

B. Spacecraft Antenna Research: High-Efficiency S- and X-Band Telemetry and Tracking Feed, K. Woo

1. Introduction

The S- and X-band telemetry and tracking feed described in SPS 37-58, Vol. III, pp. 68–72 has been modified for improved performance. This article reports the radiation characteristics and the efficiencies of the improved feed.

2. Feed Design

The design of the feed is shown in Fig. 7. As described in SPS 37-58, Vol. III, the group of four vertical probes is for monopulse tracking at 2115 MHz, the horizontal probe is for telemetering at 2295 MHz, and the vertical probe in the X-band guide is for telemetering at 8448 MHz. For the purpose of broadening and shaping the X-band pattern, a beamshaping section (in place of the step in the previous design) is introduced at the output of the X-band guide. The section consists of multiple steps in its interior. The steps generate the beamshaping modes ($\text{TE}_{12} + \text{TM}_{12}, \text{TE}_{30}$) at selected locations.

3. Feed Performance

The radiation characteristics of the feed at each operating frequency have been measured and are shown in

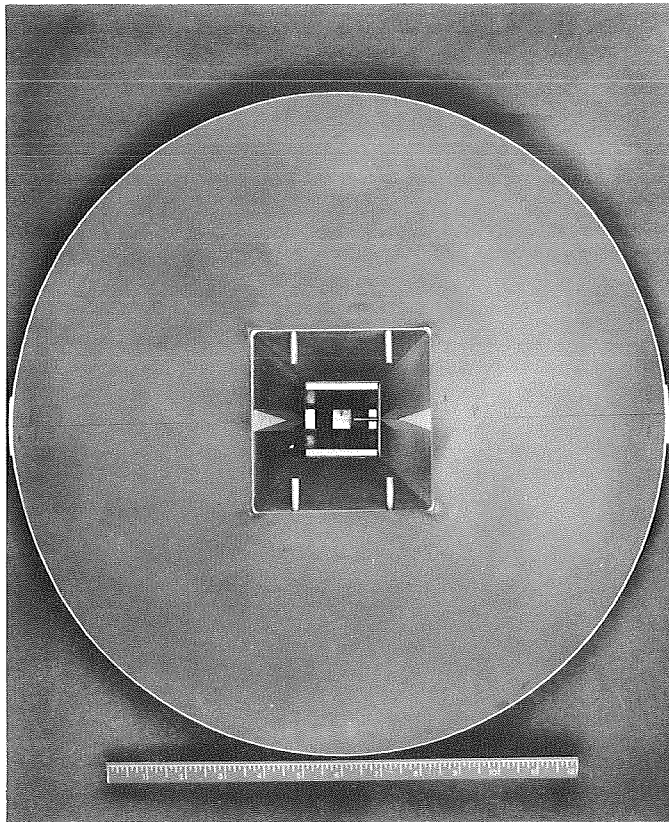


Fig. 7. Feed design

Figs. 8, 9, and 10. As can be seen, the beamwidth of the X-band pattern has been broadened to nearly those of the S-band patterns resulting high feed efficiency at both X- and S-band. The X-band pattern also provides nearly uniform illumination, which results in further enhancement of X-band efficiency. The monopulse difference channels ($\Delta V, \Delta H$) provide deep nulls on boresight as required for tracking.

The efficiency of the feed at each operating frequency has been calculated by using A. Ludwig's efficiency pro-

gram (Ref. 1). The efficiency calculation takes into account the losses due to spillover, illumination, cross-polarization, and phase. The effects of blockage and feed dissipation are yet to be determined and are therefore not included. Figure 11 shows the calculated efficiency as a function of the illumination angle. It shows that at 8448 MHz, the feed efficiency reaches a maximum of 81% at an illumination angle of 15 deg; and at 2295 and 2115 MHz, the feed efficiencies reach their respective maximums of 70% and 64% at an illumination angle of 20 deg. As can be seen, the feed achieves a very high efficiency (81%) at 8448 MHz as the result of beamshaping. Since the maximum efficiency illumination angles of X- and S-band do not coincide, the question of which illumination angle is to be used depends entirely on the efficiency requirements at these frequencies of a particular application. For example, if an illumination angle of 18 deg is chosen, the feed will provide 75% efficiency at 8448 MHz, 69% efficiency at 2295 MHz, and 62% efficiency at 2115 MHz.

4. Future Work

Future work on the feed will include:

- (1) Improving S-band efficiencies by equalizing the E- and H-plane beamwidths and achieving better phase patterns.
- (2) Further broadening the X-band beam and/or narrowing the S-band beams to bring the maximum efficiency illumination angles of X- and S-band closer together.
- (3) Improving input voltage standing-wave ratios (VSWRs) at all operating frequencies.

Reference

1. Ludwig, A., *Computer Programs for Antenna Feed System Design and Analysis*, Technical Report 32-979. Jet Propulsion Laboratory, Pasadena, Calif., Apr. 15, 1967.

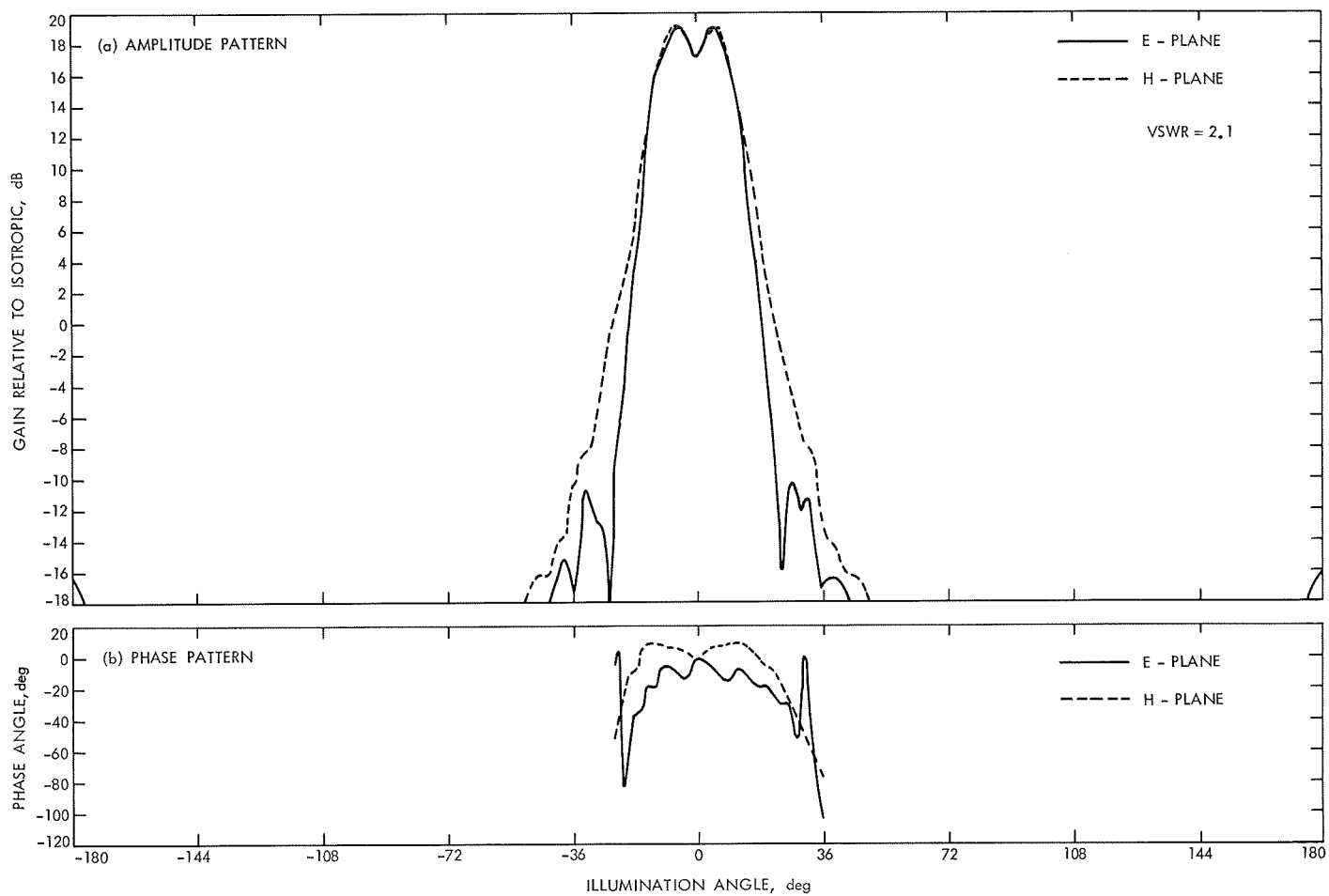


Fig. 8. Radiation characteristics of telemetry at 8448 MHz

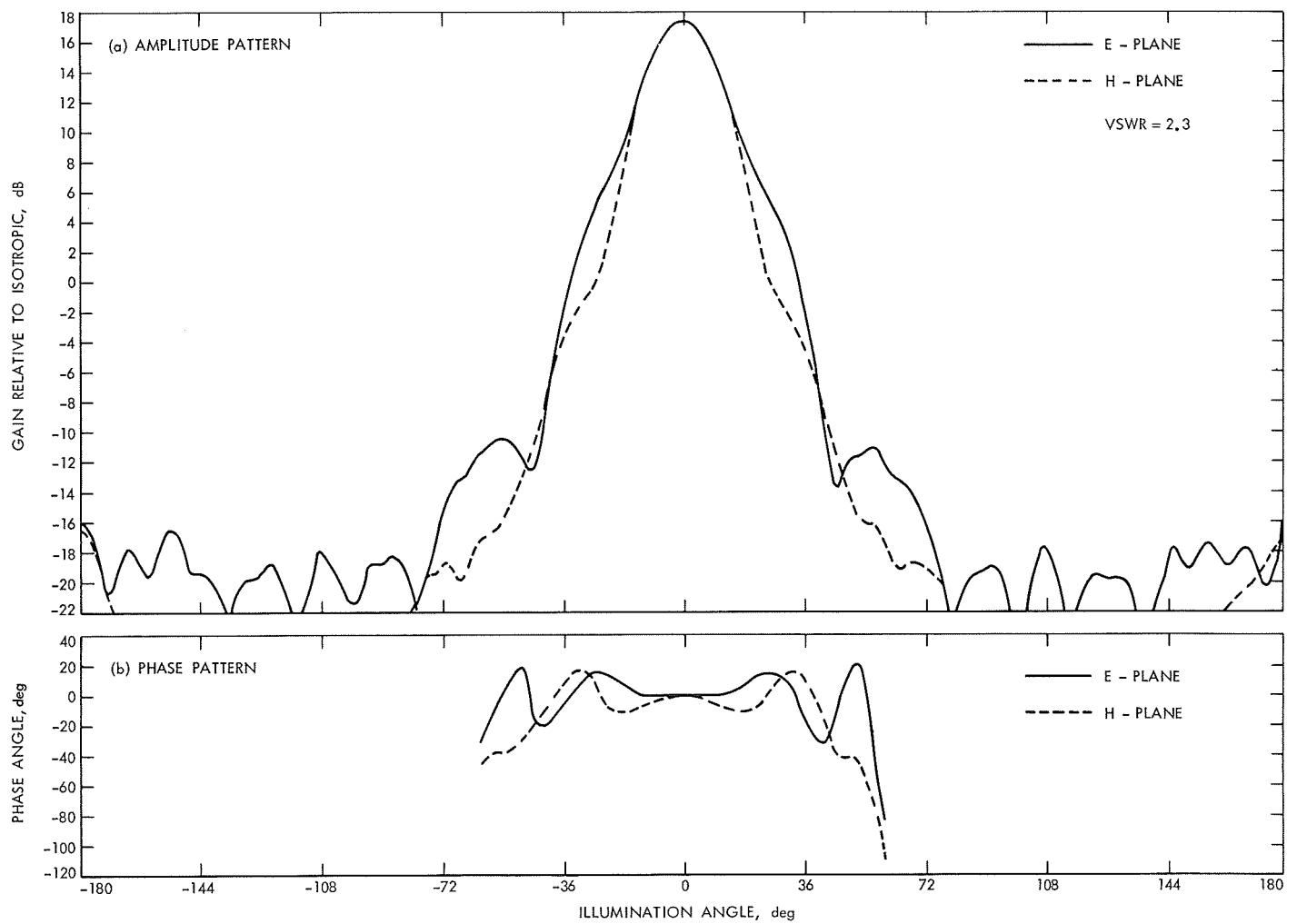


Fig. 9. Radiation characteristics of telemetry at 2295 MHz

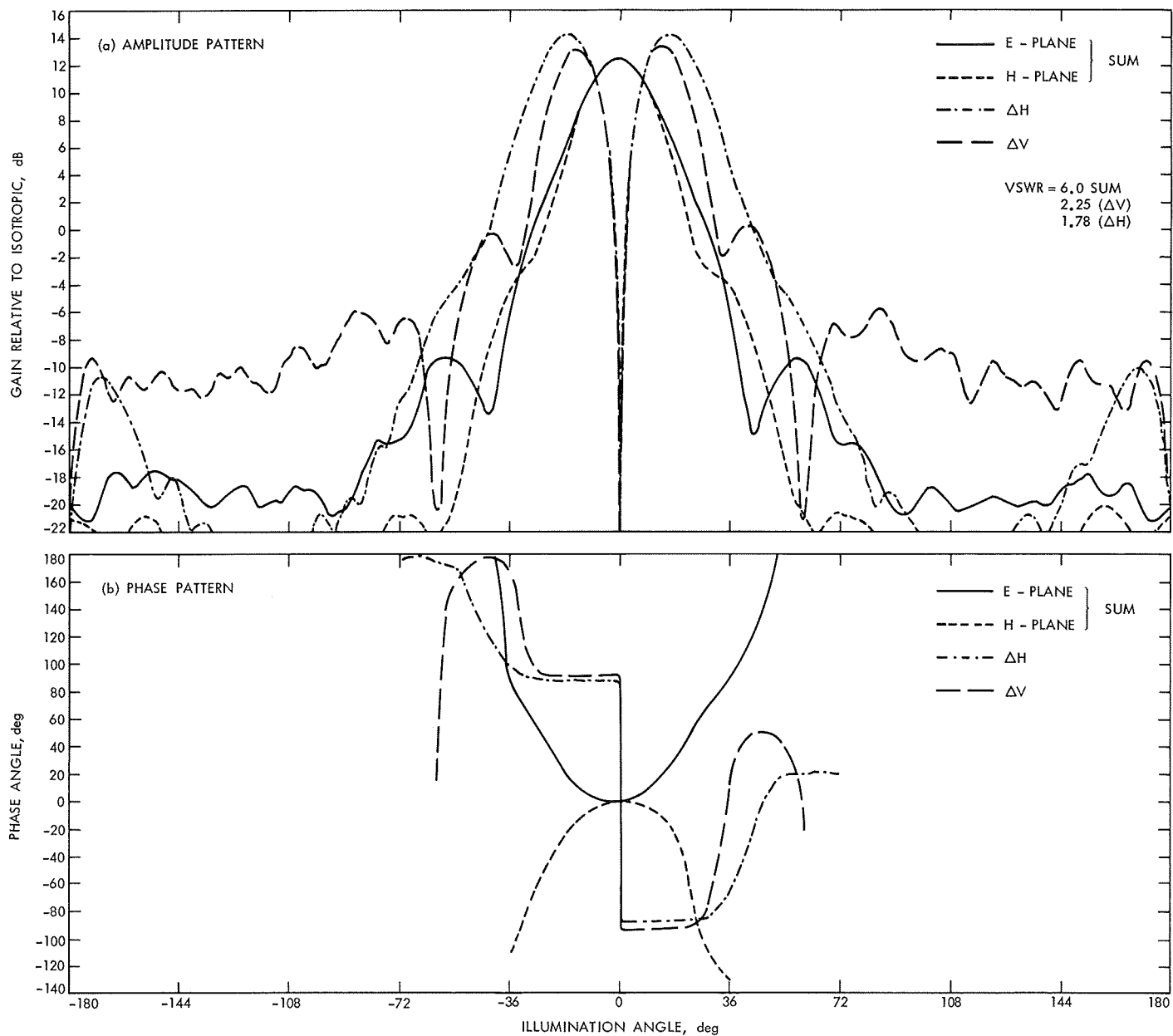


Fig. 10. Radiation characteristics of monopulse at 2115 MHz

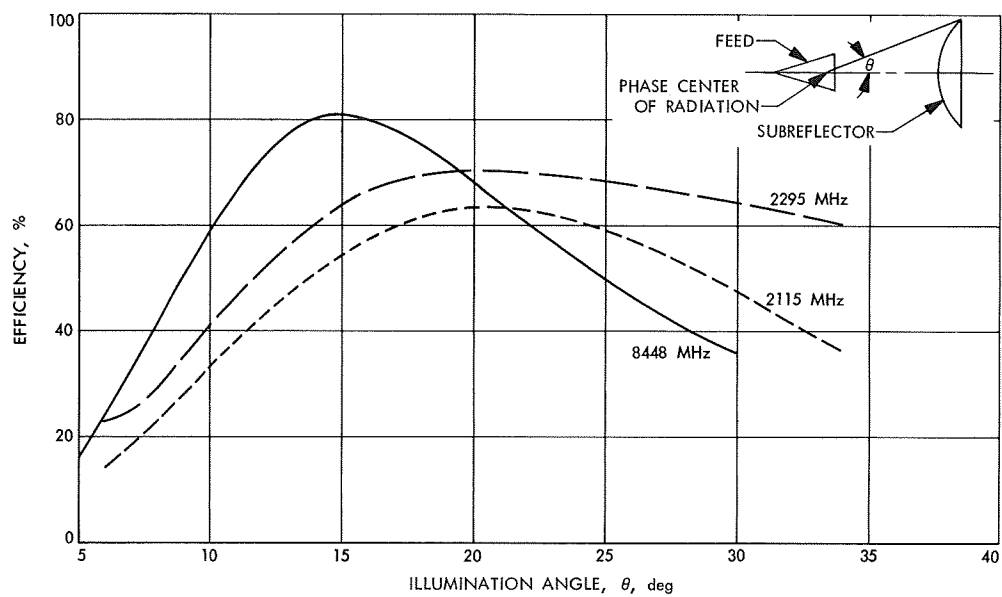


Fig. 11. Feed efficiency versus illumination angle

VII. Spacecraft Telemetry and Command

TELECOMMUNICATIONS DIVISION

A. Telemetry Data Systems Implementation,

L. A. Couvillon

1. Introduction

Planetary spacecraft have used subcarrier frequency multiplex to separate independent pulse-code-modulated (PCM) data streams. Examples are *Mariners VI* and *VII*, the 1969 missions to Mars, which transmitted engineering telemetry at a low rate ($8\frac{1}{2}$ or $33\frac{1}{2}$ bits/s) on one subcarrier, and science data at a higher rate (up to 86,400 symbols/s) on another. The two subcarriers were added in a linear summer in a ratio appropriate to the desired modulation indices, and applied to a carrier phase modulator. Turnaround ranging modulation was added in the phase modulator.

This article considers a new implementation of the frequency multiplexing and modulation functions, in order to permit programmed control of the modulation indices and consequently the ability to modify these parameters after launch. This capability is especially important to maximize the data return from a Grand Tour mission, during which the parameters of the ground receiver, and

consequently the optimum modulation indices, are likely to change significantly.

In addition to the advantage of programmability, the new technique replaces the linear summing amplifier and subcarrier modulators with digital logic. The reliability advantage of digital circuitry is therefore obtained.

2. Modulation Functions

Figure 1 is a diagram of the present technique for subcarrier modulation and combination. Each of several PCM data streams phase-shift-keyed-modulates a subcarrier. The subcarriers are then added in ratios appropriate to the desired allocation of power between subcarriers and between the subcarriers and the carrier. The subcarrier modulators and the summing amplifier have previously been implemented with linear circuitry. Provision is usually made for the selection of alternate "modes" or weights in the weighted summation. This has been accomplished by the method shown in Fig. 2 of changing the weighting resistors in an operational summing amplifier, using analog switches to add or remove impedance.

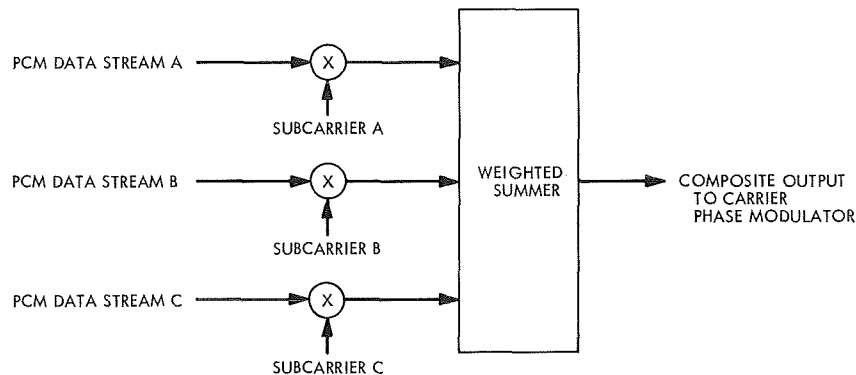


Fig. 1. Mariner Mars 1969 subcarrier modulator

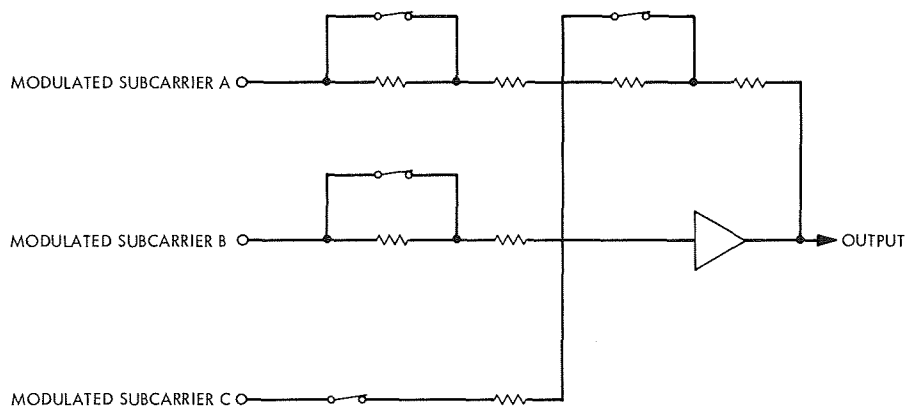


Fig. 2. Mariner Mars 1969 weighted summing amplifier

Because of the complexity of the analog switches and associated parts, the number of "modulation modes" has been constrained to a small number (three for *Mariners VI* and *VII*) and the system performance thereby compromised.

It is easily shown that under most conditions the use of square-wave signals for subcarriers results in lower intermodulation losses in a phase-modulated carrier system than with sinusoidal subcarriers. In addition, the square waves are easier to generate and manipulate in a mostly digital spacecraft. The increased bandwidth required in the detector for efficient demodulation of square waves has not been a problem in the wideband DSIF multimission telemetry system. Consequently, square-wave subcarriers are usually used. This fact results in the situation depicted in Fig. 3; the example shown considers two square-wave subcarriers of amplitude $\pm v_A$ and $\pm v_B$ added to form a *four-level* waveform. It is clear that the addition of two binary (square-wave) subcarriers always results in a four discrete level wave (or, n subcarriers leads to a 2^n level wave).

Modulation of the RF carrier with this four-level waveform results in a carrier whose phase at any instant is *one of four discrete* phases. The particular values of the discrete phases depend, of course, on the allocation of power between subcarriers and the carrier, i.e., the summing weights.

3. New Implementation

The essence of the new idea is that the discrete carrier phases can be regarded as *boolean functions of the data, subcarrier and mode control signals treated as binary variables*. Thus, if one could construct a digital carrier phase modulator (digital-to-phase converter, D/PC), the subcarrier modulators and summing amplifier are replaced by *code conversion logic* which maps the binary data, subcarriers, and control into a binary representation of phase. Figure 4 shows the data stream D_A, D_B, \dots , the subcarriers S_A, S_B, \dots , and the mode controls C_1, C_2, \dots , being transformed in "code conversion" logic to a set of phase bits which control a D/PC.

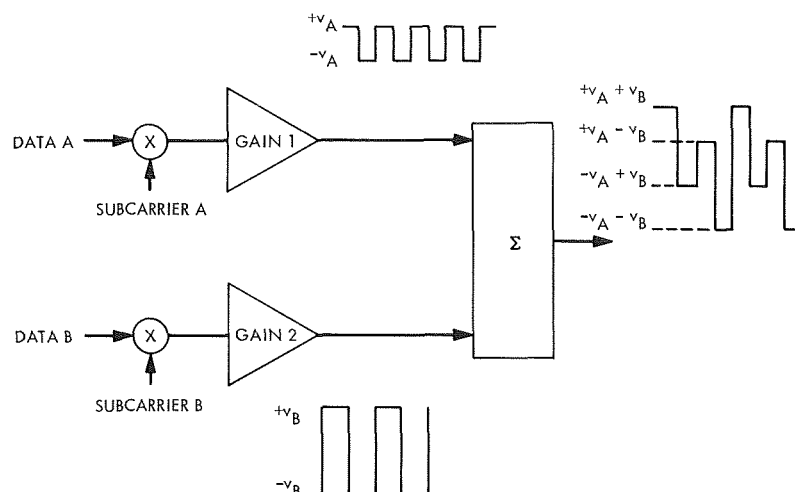


Fig. 3. Addition of square-wave subcarriers

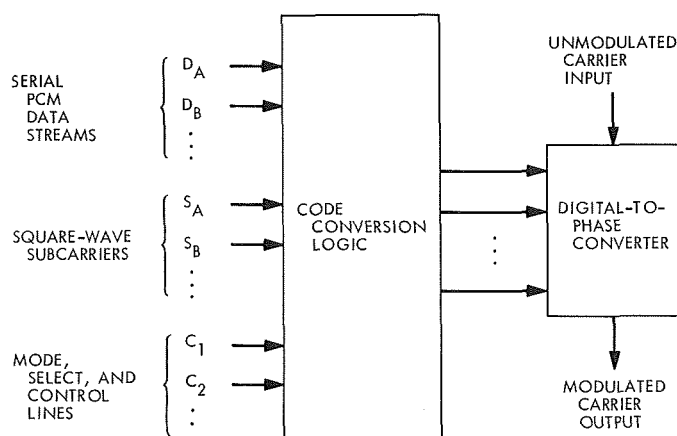


Fig. 4. Digital phase modulation

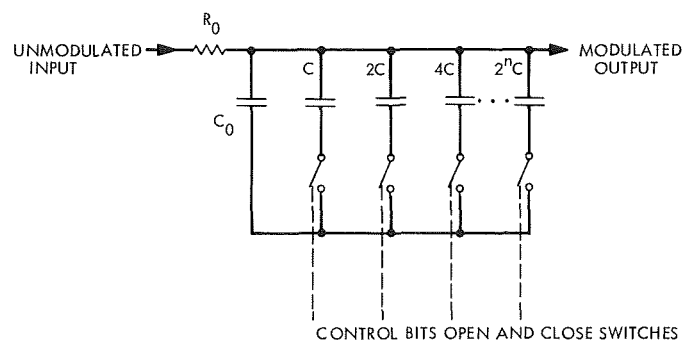


Fig. 5. Digital-to-phase converter circuit

One implementation of the D/PC circuit is shown in Fig. 5, an RC network in which the effective C is varied over a small range by the binary control. Only a small shift in RC product is necessary because the phase modulator operates at a relatively low frequency (76 MHz) and the phase deviation is multiplied during the frequency multiplication to the communication frequency of 2295 (or 8448) MHz. The phase quantization (the smallest phase increment) should be of the order of the overall system phase stability. It is felt that a seven-bit ($1/128$ resolution) D/PC is more than adequate.

The design of the code conversion logic can be accomplished in two ways. One way is to use a pair of diode

matrices—an encode matrix followed by a decode matrix—to convert the parallel k -bit number representing the data, subcarrier, and control variables into an m -bit number representing the desired carrier phase. A less tedious and more flexible method is to use a set of holding registers in which the possible carrier phase numbers are stored, transferring the appropriate register state in parallel to the D/PC through gating controlled by the data and subcarrier variables. This method is shown in Fig. 6. The in-flight programmability of this system is apparent—the holding registers are loaded, via the command system or the spacecraft computer, with the desired phase angles.

Detail design of the critical circuits, such as the D/PC, is now in progress.

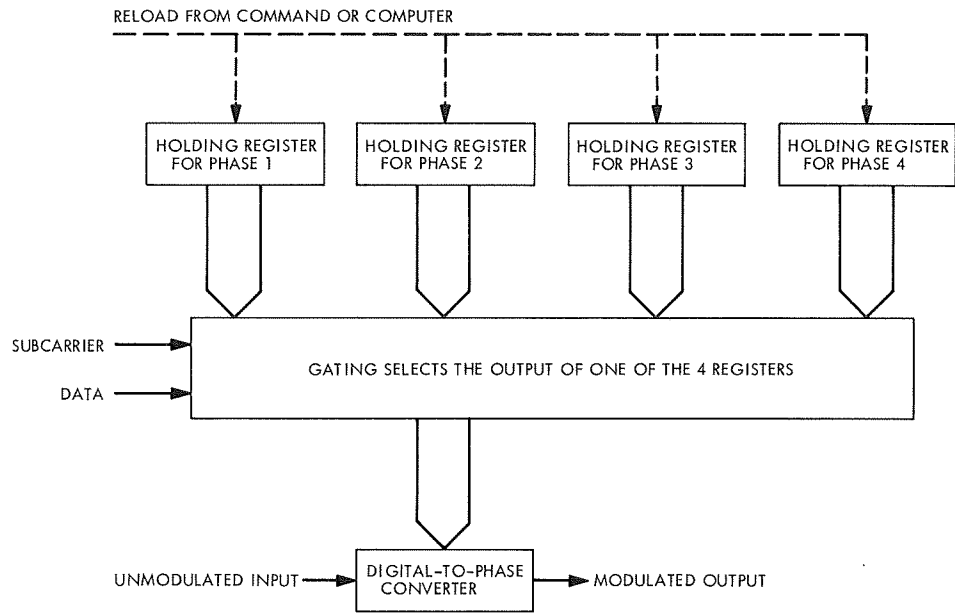


Fig. 6. Programmable modulator system

VIII. Spacecraft Telecommunications Systems

TELECOMMUNICATIONS DIVISION

A. Data-Aided Carrier Tracking Loops, W. C. Lindsey and M. K. Simon

1. Introduction

Future deep space missions are defined such that the cooperating planetary orbiting spacecraft and lander telecommunication systems are basic to the return of the increased data volume demanded by new scientific experiments, e.g., visual imaging, infrared and ultraviolet spectroscopy, and S-band occultation experiments. Consequently, there is an increased need for improved telecommunication system capability and efficiency. Aside from technological advancements, one avenue of approach toward achieving this goal employs the principles of coding theory [block-coding (Refs. 1, 2, and 3), or convolution-encoding-sequential decoding (Refs. 4, 5)]. Based upon the present state of the art in advanced coding techniques, it appears that one can achieve approximately 0.8–1.0 dB improvement (SPS 37-54, Vol. III, pp. 171–177, and SPS 37-56, Vol. III, pp. 78–83) in telemetry system efficiency above that offered by the method of block-coding employed on *Mariners VI* and *VII*. Consequently, from a theoretical point of view, it is clear at this point that one must look for new and different means for achieving a sizable improvement in system efficiency. The term system

efficiency as used here is intended to include the improvements in tracking, telemetry, and command system capability.

The purpose of this presentation is to introduce an innovation that will significantly improve the capability and performance of phase-coherent tracking, telemetry, and command systems that utilize PCM/PSK/PM modulation. The innovation should find its way into the many fields of application of telemetry and tracking, e.g., relay satellite systems, commercial television, and military telecommunication systems.

The basic idea of the innovation (Ref. 6) centers around using the power in the composite signal sidebands to enhance the effective signal-to-noise ratio (SNR) in the carrier tracking loop bandwidth, thereby reducing the noisy reference loss (Refs. 1, 2, and 7) and ultimately the probability of error of the receiver. The vehicle through which this is accomplished employs the principle of decision-directed feedback. It is simplest at the outset to discuss qualitatively the case of a received signal in the form of a carrier phase modulated by a bi-phase modulated data subcarrier, i.e., the so-called single channel system (Ref. 7). The idea here then is to form an estimate of

the bi-phase modulated data subcarrier term¹ which, when fed back to the carrier tracking loop, can be used to recover the power in the sideband components for carrier tracking purposes. The net effect is cumulative in that any reduction in the loss due to the noisy reference (Ref. 7 and footnote 1) results in a further improvement in the data detector performance (Ref. 7 and footnote 1), the subcarrier tracking loop performance,² and so on. Ordinarily, the sideband power, being centered around the subcarrier frequency, is filtered out by the carrier tracking loop (a standard phase-locked loop) and hence is not available for improving carrier tracking performance.

More specifically, consider the block diagram illustrated in Fig. 1. The upper half of the configuration constitutes a standard phase-locked loop (SPLL) used in present day tracking receivers. The modification proposed here makes use of a quadrature channel to insert a dc component into the voltage-controlled oscillator (VCO) whose power is proportional to that of the bi-phase modulated data subcarrier (Ref. 7). One principal point to note is that the shape of the equivalent S-curve is affected by the power in the signal's sidebands, the phase-jitter in the subcarrier tracking loop, the symbol-sync jitter, and the conditional probability of error of the data detector. However, cursory examination of these effects indicates that the symbol-sync jitter and subcarrier sync jitter are small relative to the radio loss caused by the carrier tracking loop phase error; consequently, essentially all of the sideband power can be

recovered and used for improving the principal source of system degradation, viz., the radio loss or noisy reference loss (Ref. 1). It appears that very little additional equipment and complexity have been introduced relative to that existing in present tracking receivers.

As an example, assuming first-order loop filters; upper and lower open loop gains, K_u and K_l , respectively; and zero initial loop detuning; we show that the loop SNR improvement, I , realized in the above data-aided carrier tracking loop relative to a SPLL is bounded from above by $I = (1 + GM)^2 / (1 + G^2)$. Here $G = K_u / K_l$ and

$$M = [(1 - m^2) / m^2]^{1/2}$$

where m is the modulation factor (Ref. 7). Moreover, the ratio of the bandwidths of the data-aided PLL to that of the SPLL is given by

$$W_{Le} / W_L = (K_u / K_l) (1 + GM)$$

where K is the open loop gain of the SPLL. By selecting the ratio G , say G_0 , such that I is maximized subject to the constraint $W_{Le} / W_L = 1$, one finds that

$$G_0 = M$$

$$I = 1 / m^2$$

$$K_u = m^2 K$$

$$K_l = [(1 - m^2) / m^2]^{1/2} K$$

Presently, the current state of the art demands that $m^2 \geq 0.1$; hence the maximum improvement in SNR performance for one-way tracking systems is, under the above assumptions, $I = 10$ dB. The corresponding enhancement

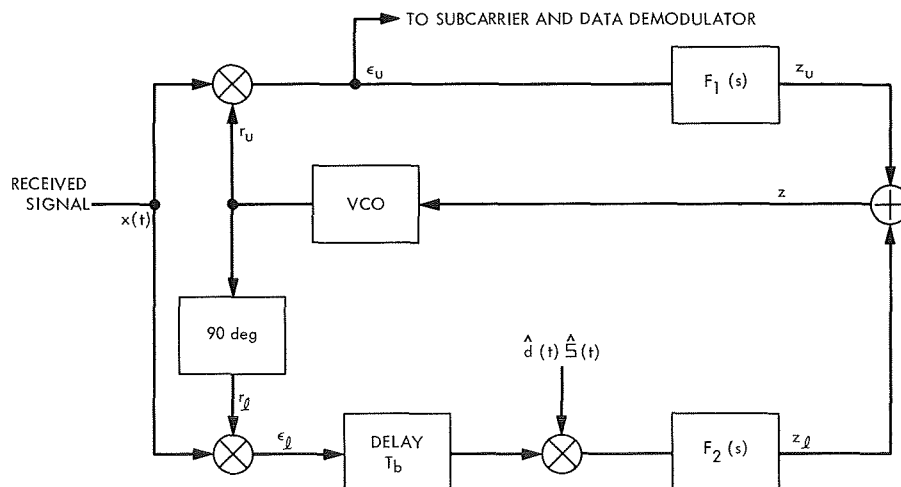


Fig. 1. Data-aided carrier tracking loop

in telemetry and command system performance is difficult to assess without carrying out a tedious but tractable analytical study. However, it is known that the maximum relative improvement in the communication efficiency factor E_b/N_0 (energy per bit/noise spectral density) is

$$I = (\overline{\cos \phi_2})^2 / (\overline{\cos \phi_1})^2$$

where $(\overline{\cos \phi_2})^2$ and $(\overline{\cos \phi_1})^2$ are, respectively, the normalized radio losses suffered in the data channel using the technique described herein and as of the current state of the art (Ref. 8).

In situations where several data and/or sync modulated subcarriers are phase modulated on to a main carrier (Ref. 3), an n -dimensional extension of this technique can be applied. The power in any of the subcarrier components can, in principle, be recovered as above. Furthermore, cross-modulation losses can be recovered and used for carrier tracking purposes as follows. Several decision-directed subcarrier data demodulators are employed in parallel. Each of these requires an input delay element proportional to the reciprocal of the data rate in that channel. In order to form an estimate of the cross-modulation component of interest, the outputs of the channels involved must be delayed by the complementary delay before multiplication. The complement is taken with respect to the least common period of the channel data rates. Sufficient storage must also be provided in each channel for storing the bit decisions made during the complementary delay. The delay required in the carrier tracking loop is then equal to this least common period. An example of this extension (typical of *Mariners VI* and *VII* systems) is illustrated in Fig. 2 for a carrier

phase modulated by two modulated data subcarriers. The n -dimensional extension of this concept is obvious. We now explore the quantitative aspects of data-aided carrier tracking loops.

2. Development of the Stochastic Differential Equation of Operation

In this subsection, we develop the loop equation of operation to the extent that it applies to Fig. 1. It is then a trivial matter to modify this equation for application to the system illustrated in Fig. 2, which can be used for recovering cross-modulation losses. Before proceeding, it should be noted that we will draw heavily upon the notation introduced in Refs. 7 and 9. For a single-channel communication system, the received signal $x(t)$ is of the form (Ref. 7)

$$\begin{aligned} x(t) &= (2P)^{1/2} \sin [\omega_0 t + (\cos^{-1} m) x(t) + \theta] + n_i(t) \\ &= (2m^2 P)^{1/2} \sin \Phi(t) \\ &\quad + [2(1 - m^2) P]^{1/2} x(t) \cos \Phi(t) + n_i(t) \end{aligned} \quad (1)$$

where P is the average radiated power, m^2 is the modulation factor, $x(t) = d(t) \ominus(t)$ is a bi-phase modulated data subcarrier, and θ is a slowly varying function of time related to the Doppler offset. The function $\Phi(t)$ is defined by $\Phi(t) = \omega_0 t + \theta$ and the random process $\{n_i(t)\}$ is assumed to be narrowband and Gaussian. A sample function representation of this process is given by (Ref. 10)

$$n_i(t) = 2^{1/2} [n_1(t) \cos \omega_0 t - n_2(t) \sin \omega_0 t] \quad (2)$$

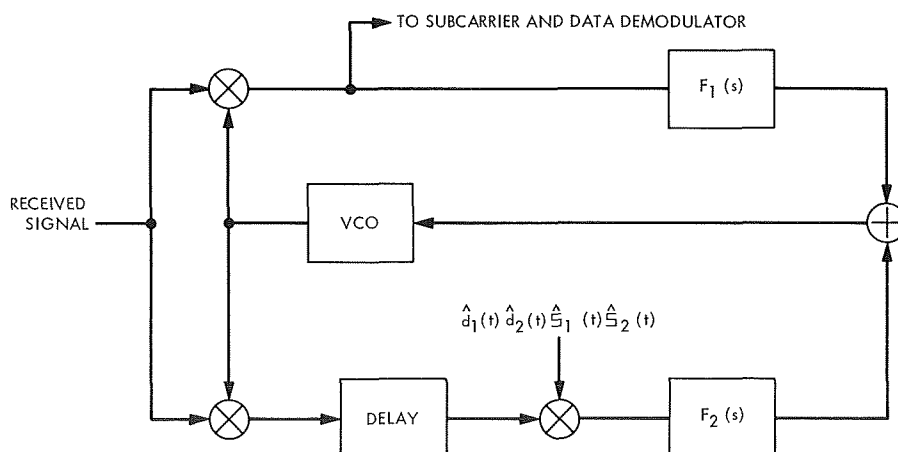


Fig. 2. Cross-modulation-aided carrier tracking loop

where $n_1(t)$ and $n_2(t)$ are independent Gaussian random processes with single-sided spectral densities, N_0 W/Hz. It is convenient to characterize the reference signals $r_u(t)$ and $r_i(t)$ in Fig. 1 by

$$\begin{cases} r_u(t) = 2^{1/2} K_1 \cos \hat{\Phi}(t) \\ r_i(t) = -2^{1/2} K_2 \cos \hat{\Phi}(t) \end{cases} \quad (3)$$

where $\hat{\Phi}(t) = \omega_0 t + \hat{\theta}$, and $\hat{\theta}$ is the loop estimate of θ . Assuming an upper loop multiplier gain K_{1m} and neglecting double frequency terms, it is easy to show that, in operator form,

$$z_u(t) = K_1 K_{1m} F_1(p) [(m^2 P)^{1/2} \sin \phi(t) + n_u(t)] \quad (4)$$

where $n_u(t)$ is a baseband white Gaussian process (Ref. 11) of N_0 W/Hz, and $\phi(t) = \Phi(t) - \hat{\Phi}(t)$ is the loop phase error.

Similarly, assuming a lower loop multiplier gain of K_{2m} and neglecting double frequency components, the output signal z_i is given by

$$\begin{aligned} z_i(t) &= K_2 K_{2m} F_2(p) \exp(-pT_b) \\ &\times [S^{1/2} d(t) \hat{d}(t) \hat{S}(t) \sin \phi(t) + n_i(t)] \end{aligned} \quad (5)$$

where $\hat{d}(t)$ is the estimate of the data $d(t)$, $\hat{S}(t)$ is the estimate of the subcarrier reference, $S = (1 - m^2)P$ is the sideband power, and T_b is the time per bit (symbol). In arriving at Eq. (5), we have also assumed that a bandpass filter removes from the dynamic phase error ϵ_i the bi-phase modulated data subcarrier. Consequently, $n_i(t)$ can be conveniently modeled as a bandlimited white Gaussian noise process of N_0 W/Hz. We point out that $n_i(t)$ and $n_u(t)$ are statistically independent random processes. Hereafter, for simplicity of notation, we omit the dependence on t in all variables that are functions of time.

When the bandwidth of the loop is small relative to the signalling data rate, $\mathcal{R}_b = T_b^{-1}$, then on the average Eq. (5) reduces to³

$$z_i = K_2 K_{2m} F_2(p) \{S^{1/2} [1 - 2P_E(\phi)] \sin \phi + n_i\} \quad (6)$$

³The transfer function $\exp(-pT_b)$ with $p = j\omega$ is approximately unity for all ω of interest when $W_{Le} T_b \ll 1$ where W_{Le} is the equivalent bandwidth of the data-aided loop to be subsequently defined.

where for the special case of phase-shift keyed signals (Ref. 7),

$$P_E(\phi) = \frac{1}{(2\pi)^{1/2}} \int_{(2R)^{1/2} \cos \phi}^{\infty} \exp(-x^2/2) dx \quad (7)$$

is the symbol error probability, and $R = ST_b/N_0$.

Making the usual assumption that the instantaneous VCO phase $\hat{\theta}$ is related to its input z through

$$\hat{\theta} = \frac{K_v}{p} z = \frac{K_v}{p} [F_1(p)(x r_u) + F_2(p)(x r_i) \hat{d} \hat{S}] \quad (8)$$

where K_v is the gain of the VCO and since $z = z_u + z_i$ and $\phi = \Phi - \hat{\Phi}$, we find from Eqs. (4), (6) and (8) and the fact that $\phi = \theta - \hat{\theta}$,

$$\begin{aligned} \dot{\phi} &= \dot{\theta} - K_u F_1(p) [(m^2 P)^{1/2} \sin \phi + n_u] \\ &\quad - K_i F_2(p) \{S^{1/2} [1 - 2P_E(\phi)] \sin \phi + n_i\} \end{aligned} \quad (9)$$

where $K_u = K_1 K_{1m} K_v$ and $K_i = K_2 K_{2m} K_v$, respectively, represent the upper and lower open loop gains. Equation (9) represents the stochastic differential equation of loop operation for the configuration illustrated in Fig. 1.

3. Nonlinear Analysis of First-Order Loops

In this section, we present the analysis for first-order loops. Based upon the work accomplished in Ref. 9, extension of these results to higher-order loops will be forthcoming. However, the work that follows is certainly indicative of the capability that might be expected from higher-order loops. For the case where the loop filters $F_1(p)$ and $F_2(p)$ have unit gains, Eq. (9) reduces to

$$\dot{\phi} = \Omega_0 - \{K_u (m^2 P)^{1/2} + K_i S^{1/2} [1 - 2P_E(\phi)]\} \sin \phi + n_e(t) \quad (10)$$

where we have assumed that $\theta = \Omega_0 + \theta_0$ and $n_e(t) = n_u(t) + n_i(t)$ is a white Gaussian noise process with spectral density

$$S_{n_e}(\omega) = \frac{N_0 K_u^2}{2} (1 + G^2) \quad (11)$$

with $G \triangleq K_i/K_u$.

a. Steady-state phase-error probability density function. Using results given in Ref. 9, it is easy to show that

the probability density function (PDF) of the phase-error process is given by

$$p(\phi) = N \exp [U_0(\phi)] \int_{\phi}^{\phi+2\pi} \exp [-U_0(x)] dx; |\phi| \leq \pi \quad (12)$$

where N is the normalization constant and the potential function $U_0(\phi)$ is given by⁴

$$U_0(\phi) = \frac{2}{K_{00}} \left\{ \Omega_0 \phi + K_u (m^2 P)^{1/2} \cos \phi + K_t S^{1/2} \left[\cos \phi \operatorname{erf}(R^{1/2} \cos \phi) + \frac{1}{(\pi R)^{1/2}} \exp(-R \cos^2 \phi) \right] \right\} \quad (13)$$

Furthermore, the diffusion coefficient, K_{00} , is given by

$$K_{00} = \frac{N_0 K_u^2}{2} (1 + G^2) \quad (14)$$

and $\operatorname{erf}(x)$ is defined by

$$\operatorname{erf}(x) = \frac{2}{\pi^{1/2}} \int_0^x \exp(-y^2) dy \quad (15)$$

To optimize the design, one desires to minimize the variance of the phase error for a given m by proper choice of the loop gains K_u and K_t . This optimization procedure is extremely tedious and must be carried out by means of numerical integration, etc., on a digital computer. Consequently, at this point we propose to optimize the loop as follows.

Let A, K be taken as the parameters of a SPLL (Ref. 9). Then the SNR in the loop bandwidth is given by (Ref. 9)

$$\rho = 4A/N_0 K = 2A^2/N_0 W_L$$

where $W_L = AK/2$ is the two-sided loop bandwidth. For the data-aided carrier tracking loop, there exists, in its linear region of operation, an analogous parameter

$$\rho_e = \frac{4A_e}{N_0 K_e} = \frac{2A_e^2}{N_0 W_{Le}} \quad (16)$$

which we shall refer to as the effective loop SNR of the data-aided carrier tracking loop. Thus, in terms of the

⁴For the remainder of this paragraph, we pursue the phase-shift-keyed (PSK) signal case wherein $P_E(\phi)$ is given by Eq. (7).

parameters of a SPLL,

$$A_e = A \left[\frac{1 + GM \operatorname{erf} R^{1/2}}{(1 + G^2)^{1/2}} \right] \quad (17)$$

where again M is defined by $M = [(1 - m^2)/m^2]^{1/2}$ and

$$K_e = K \left(\frac{K_u}{K} \right) [(1 + G^2)^{1/2}] \quad (18)$$

Using Eqs. (16), (17), and (18) and the definition for ρ we find that

$$I \triangleq \frac{\rho_e}{\rho} = \left(\frac{K}{K_u} \right) \left(\frac{1 + GM \operatorname{erf} R^{1/2}}{1 + G^2} \right) \quad (19)$$

and

$$\frac{W_{Le}}{W_L} = \left(\frac{K_u}{K} \right) (1 + GM \operatorname{erf} R^{1/2}) \quad (20)$$

The optimization procedure is carried out as follows: Choose K_u and G such that ρ_e is maximized subject to the constraint that $W_{Le}/W_L = 1$. Since $\rho_e = 2A_e^2/N_0 W_{Le}$, we equivalently wish to maximize A_e^2 with respect to G .

Differentiating Eq. (17) with respect to G and equating to zero yields an optimum value of G , say G_0 , given by

$$G_0 = M \operatorname{erf} R^{1/2} = \left(\frac{1 - m^2}{m^2} \right)^{1/2} \operatorname{erf} R^{1/2} \quad (21)$$

Also if $W_{Le}/W_L = 1$ then

$$\frac{K}{K_u} = 1 + G_0 M \operatorname{erf} R^{1/2} \quad (22)$$

so that the maximum value of I , say I_{\max} , is given by

$$I_{\max} \triangleq \left(\frac{\rho_e}{\rho} \right)_{\max} = \frac{(1 + G_0 M \operatorname{erf} R^{1/2})^2}{1 + G_0^2} \quad (23)$$

Substitution for G_0 into Eq. (23) gives

$$I_{\max} = \frac{1}{m^2} [\operatorname{erf}^2 R^{1/2} + m^2 (1 - \operatorname{erf}^2 R^{1/2})] \quad (24)$$

and

$$\left. \begin{aligned} K_u &= \left[\frac{m^2}{\operatorname{erf}^2 R^{1/2} + m^2 (1 - \operatorname{erf}^2 R^{1/2})} \right] K \\ K_t &= \left(\frac{1 - m^2}{m^2} \right)^{1/2} \left[\frac{m^2 \operatorname{erf} R^{1/2}}{\operatorname{erf}^2 R^{1/2} + m^2 (1 - \operatorname{erf}^2 R^{1/2})} \right] K \end{aligned} \right\} \quad (25)$$

As mentioned in the introduction, for $m^2 = 0.1$ and $R \rightarrow \infty$ the relative improvement becomes $I_{\max} = 10$ dB.

For equal bandwidths, i.e., $W_{Le}/W_L = 1$, the potential function Eq. (13) becomes

$$U_0(\phi) = \frac{2\Omega_0}{K_{00}} \phi + \frac{\rho(1 + GM \operatorname{erf} R^{1/2})}{1 + G^2} \times \left\{ [1 + GM \operatorname{erf}(R^{1/2} \cos \phi)] \cos \phi + \frac{GM}{(\pi R)^{1/2}} \exp(-R \cos^2 \phi) \right\} \quad (26)$$

Substitution for G_0 from Eq. (21) gives

$$U_0(\phi) = \frac{2\Omega_0}{K_{00}} \phi + \rho \left\{ [1 + M^2 \operatorname{erf}^2(R^{1/2} \cos \phi)] \cos \phi + \frac{M^2}{(\pi R)^{1/2}} \operatorname{erf}(R^{1/2} \cos \phi) \exp(-R \cos^2 \phi) \right\} \quad (27)$$

The variance of the phase error is then given by

$$\sigma_{\phi_e}^2 = \int_{-\pi}^{\pi} \phi^2 p(\phi) d\phi \quad (28a)$$

and must be evaluated numerically on a digital computer using Eqs. (12), (27) and (28). For values of m^2 where $p(\phi)$ becomes trimodal within the interval $(-\pi, \pi)$, it may be desirable to apply an alternate definition for the variance of the phase error. The reason for this is focused on the fact that the system now possesses two additional unstable singularities (i.e., ϕ_1, ϕ_2) in the interval $(-\pi, \pi)$, and hence the phase values $-\pi, \pi$ act as stable nodes. If one is able to mechanize the system so as to resolve the ambiguity between the stable points 0 and π (or $-\pi$), then the variance of the phase error should be defined as

$$\sigma_{\phi_e}^2 = \int_{-\pi}^{-\phi_1} (\pi + \phi)^2 p(\phi) d\phi + \int_{-\phi_1}^{\phi_2} \phi^2 p(\phi) d\phi + \int_{\phi_2}^{\pi} (\pi - \phi)^2 p(\phi) d\phi \quad (28b)$$

where ϕ_1 and ϕ_2 are the values of the phase error for which the potential function $U_0(\phi)$ has its minimum in an interval of width 2π . It is a simple matter to show that for the case of zero detuning, $p(\phi)$ is symmetric in ϕ and $\phi_1 = \phi_2$, thus allowing simplification of Eq. (28b).

b. Mean time to first loss of phase synchronization, $\tau_e(2\pi)$. Based upon the theory presented in Ref. 9, it is easy to show that the equivalent bandwidth-first slip time product is given by

$$W_{Le} \tau_e(2\pi) = \frac{\rho_e}{2} \int_0^\pi \frac{dx}{p(x)} \quad (29)$$

where $p(x)$ is given by Eqs. (12) and (26) and $\Omega_0 = 0$ is assumed. When $G = G_0$, $W_{Le}/W_L = 1$, Eq. (29) reduces to

$$W_{Le} \tau_e(2\pi) = \frac{\rho}{2m^2} [\operatorname{erf}^2 R^{1/2} + m^2 (1 - \operatorname{erf}^2 R^{1/2})] \int_0^\pi \frac{dx}{p(x)} \quad (30)$$

For the SPLL (Refs. 9 and 11)

$$W_L \tau(2\pi) = \pi^2 \rho I_0^2(\rho) \quad (31)$$

so that

$$\frac{\tau_e(2\pi)}{\tau(2\pi)} = \frac{\operatorname{erf}^2 R^{1/2} + m^2 (1 - \operatorname{erf}^2 R^{1/2})}{2\pi^2 m^2 I_0^2(\rho)} \int_0^\pi \frac{dx}{p(x)} \quad (32)$$

where $p(x)$ is determined from Eqs. (12) and (27).

For large ρ , the relative improvement in mean slip time-bandwidth product becomes

$$\frac{W_{Le} \tau_e(2\pi)}{W_L \tau(2\pi)} \simeq \exp[2(\rho_e - \rho)] \quad (33)$$

while for $G = G_0$, and $W_{Le}/W_L = 1$,

$$\frac{\tau_e(2\pi)}{\tau(2\pi)} \simeq \exp \left\{ 2\rho \left[\frac{(1 - m^2) \operatorname{erf}^2 R^{1/2}}{m^2} \right] \right\} \quad (34)$$

Finally, for small ρ , $I_0(\rho) \approx 1$, $p(\phi) \approx 1/2\pi$ and Eq. (33) reduces to

$$\frac{\tau_e(2\pi)}{\tau(2\pi)} \approx \frac{1}{m^2} [\operatorname{erf}^2 R^{1/2} + m^2 (1 - \operatorname{erf}^2 R^{1/2})] \quad (35)$$

c. System performance and comparisons with a SPLL. Before we consider detailed performance of the loop, it is of interest to study the equivalent normalized loop S-curve with $\Omega_0 = 0$ as given by

$$\mathcal{S}(\phi) = \frac{A_e g(\phi)}{(m^2 P)^{1/2}} = \frac{\{1 + GM[1 - 2P_E(\phi)]\}}{(1 + G^2)^{1/2}} \sin \phi \quad (36)$$

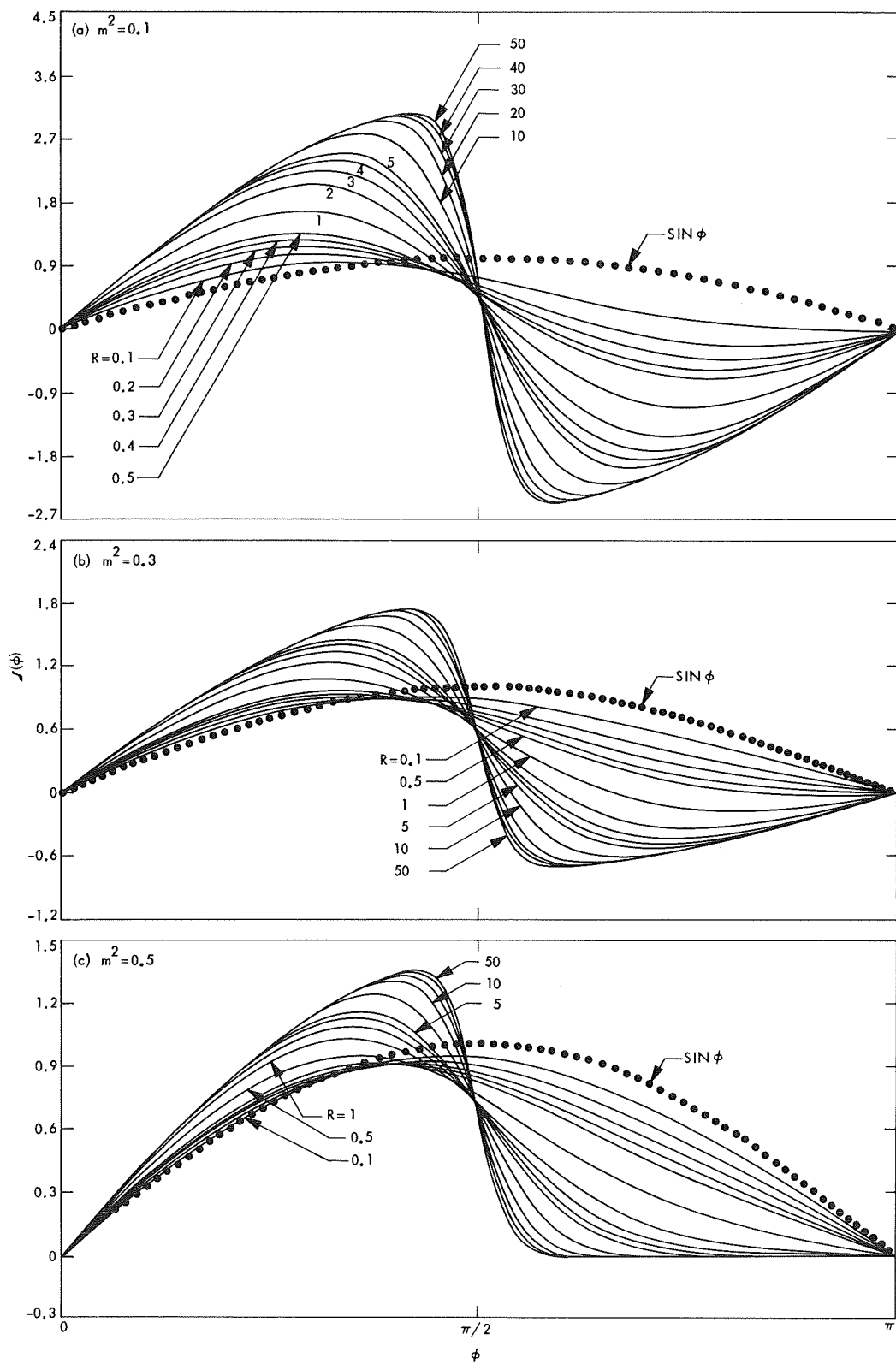


Fig. 3. Normalized loop S-curve versus phase error with $R = ST_b/N_0$ as a parameter

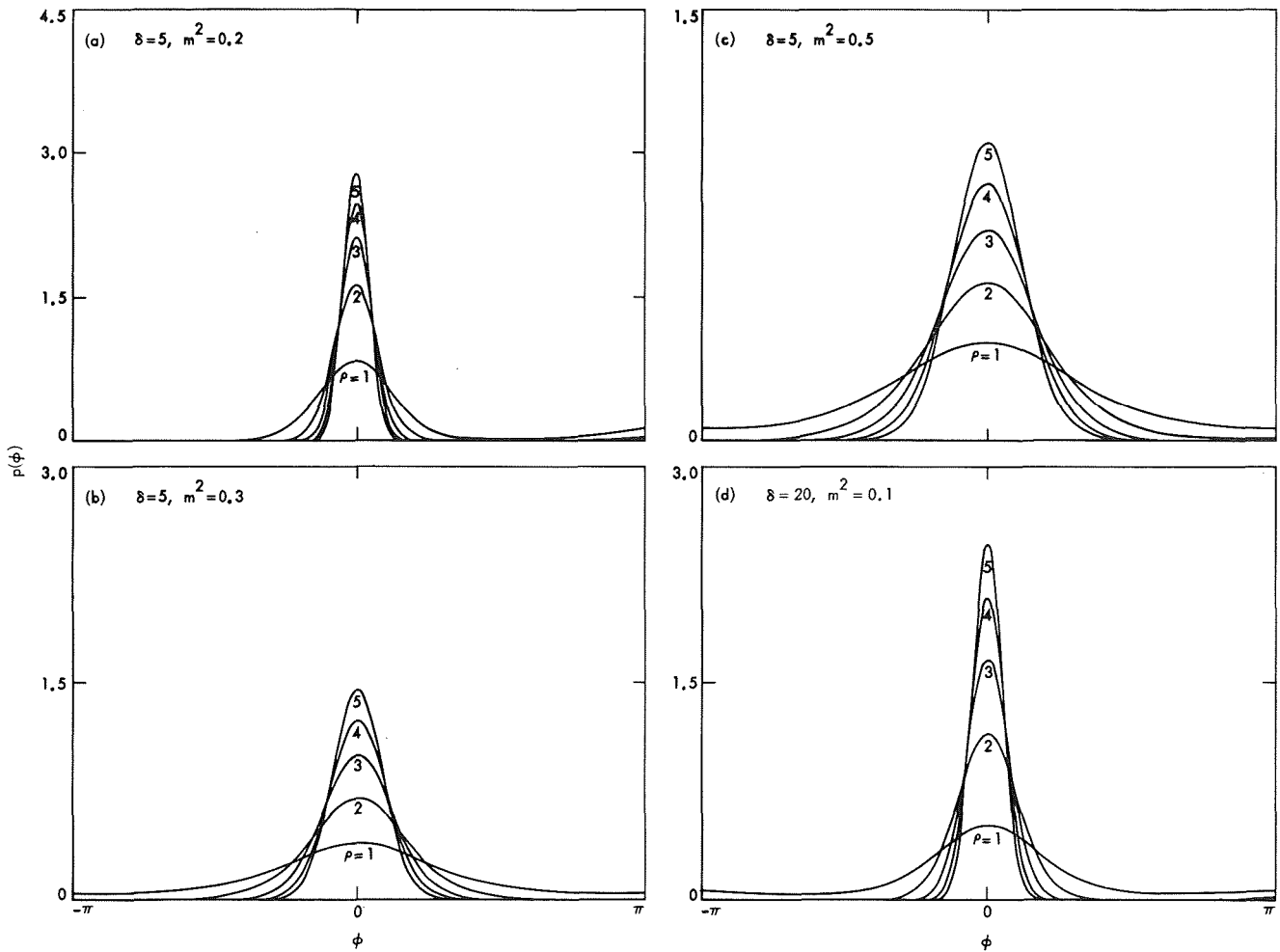


Fig. 4. Phase error PDF versus phase error with loop SNR of SPLL as a parameter ($\rho \geq 1$)

where $g(\phi)$ is the equivalent loop nonlinearity of a generalized tracker. With $G = G_0$ and $W_{Le}/W_L = 1$,

$$\mathcal{S}(\phi) = \frac{1 + M^2 [1 - 2P_E(\phi)] [1 - 2P_E(0)]}{\{1 + M^2 [1 - 2P_E(0)]^2\}^{1/2}} \sin \phi \quad (37)$$

Illustrated in Fig. 3 are the normalized loop S-curves with m^2 and $\delta = \mathcal{R}_b/W_L = (W_L T_b)^{-1}$ as parameters, and $P_E(\phi)$ as given by Eq. (7). For comparison purposes, a dotted line has been included to show the S-curve of a SPLL. This S-curve is obtained from Eq. (36) by letting $G = 0$, i.e., opening the lower loop.

For $\Omega_0 = 0$, the PDF, $p(\phi)$, given by Eqs. (12) and (27) (i.e., with $G = G_0$ and $W_{Le}/W_L = 1$) is illustrated in Figs. 4 and 5 for various values of ρ with δ and m^2 as parameters. Figures 4a and 4d can be used to study how

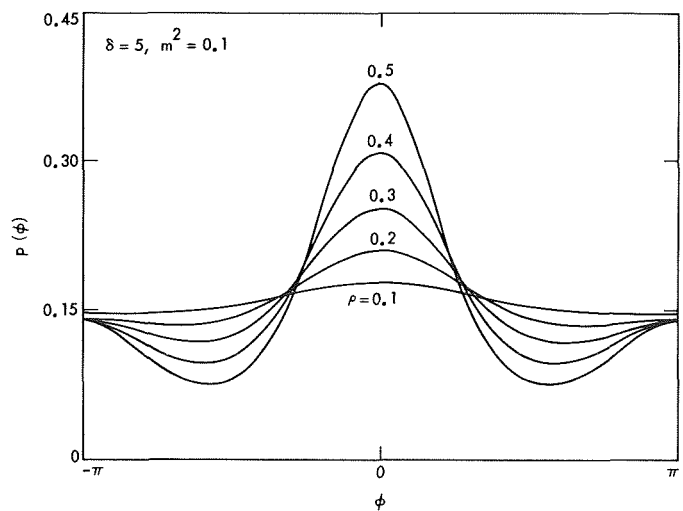


Fig. 5. Phase error PDF versus phase error with loop SNR of SPLL as a parameter ($\rho < 1$)

$p(\phi)$ is affected by a change in δ . It is interesting to note from Fig. 5 that for $\rho < 1$, the PDF becomes tri-modal.

The mean slip time, $\tau_e(2\pi)$, as computed from Eqs. (12), (27), and (30), for the data-aided carrier tracking loop is plotted in Fig. 6 versus ρ for fixed δ and several values of m^2 . It is important to note the increase in slope of these curves with decreasing m^2 in accordance with Eq. (30). Illustrated in Fig. 7 is the behavior of the variance of the phase error, as computed from Eqs. (12), (27), and (28a) or (28b), as a function of ρ for various values of m^2 with δ fixed. For the case $m^2 = 0.5$, the results, as predicted by Eqs. (28a) and (28b), are identical; hence, only one curve

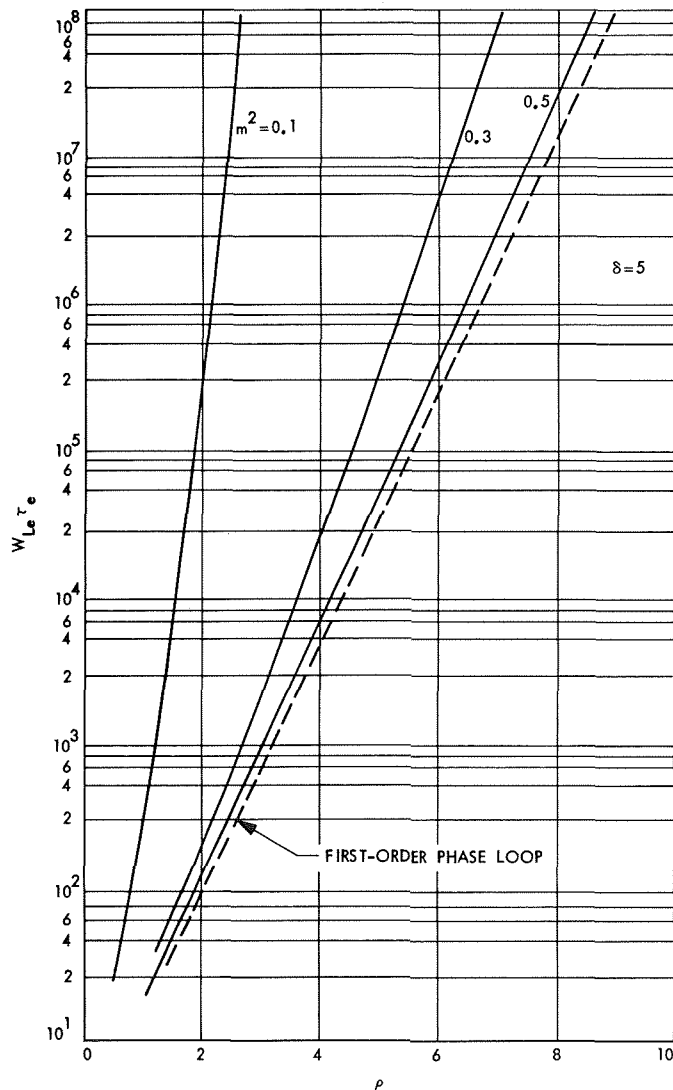


Fig. 6. Mean time to first slip versus loop SNR of SPLL with modulation factor as a parameter

appears. Figure 8 represents an alternate presentation of the results of Fig. 7 by comparing the performance of the data-aided carrier-tracking loop to that of a SPLL. More specifically, this curve depicts the ratio of the mean square phase error in the data-aided loop to that of a SPLL (in dB) versus ρ for various values of m^2 with $\delta = 5$. Again note that the limiting asymptotes as predicted by Eq. (24) with $R \rightarrow \infty$ are indicated by dashed lines on the figure. This is so since in the limit as $\rho \rightarrow \infty$, $\sigma_\phi^2/\sigma_{\phi e}^2 \approx \rho_e/\rho$. It is interesting to note that one can, in principle, obtain a relative improvement in mean square phase error in the

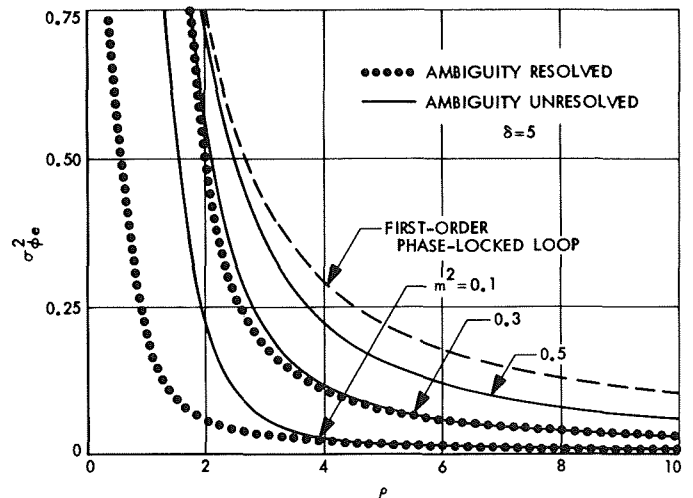


Fig. 7. Variance of the phase error versus loop SNR of SPLL with modulation factor as a parameter

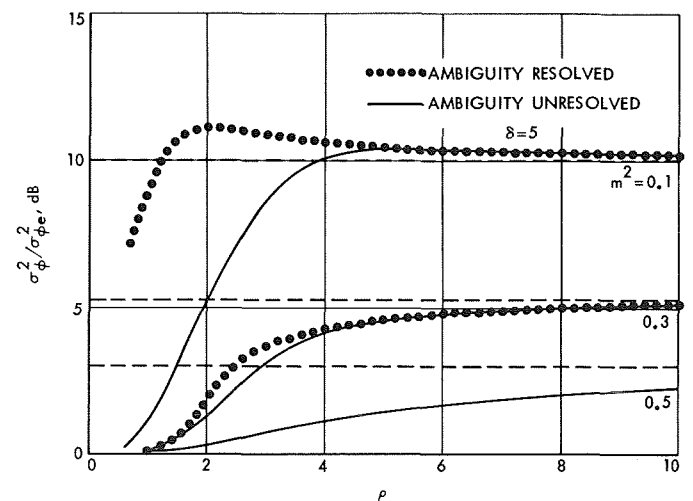


Fig. 8. Relative improvement (in dB) of the mean square phase error of the data-aided loop over that of the SPLL with modulation factor as a parameter

nonlinear region of operation greater than that predicted for linear behavior. This property is undoubtedly inherent in the nonlinear nature of the loop S-curves.

To illustrate the equivalent performance of a coherent data demodulator, we make use of the conditional error probability as given in Ref. 7 and Eq. (7) to write

$$P_E = \int_{-\pi}^{\pi} p(\phi) P_E(\phi) d\phi \quad (38a)$$

or

$$P_E = \int_{-\pi}^{-\phi_1} p(\phi) P_E(\pi + \phi) d\phi + \int_{-\phi_1}^{\phi_2} p(\phi) P_E(\phi) d\phi + \int_{\phi_2}^{\pi} p(\phi) P_E(\pi - \phi) d\phi \quad (38b)$$

where $p(\phi)$ is determined from Eqs. (12) and (27) and ϕ_1, ϕ_2 have previously been defined. The application of Eqs. (38a) or (38b) to a particular situation once again depends upon the ability, in practice, to resolve the ambiguities between $-\pi$ and π when, in fact, they occur. Performing the numerical integration necessary for evaluation of Eqs. (38a) and (38b) produces the results given in Fig. 9. For comparison purposes, we superimpose (1) a result from Ref. 5 for the case where the power between the carrier and data sidebands is divided optimally, and (2) the ideal performance curve. It should be reiterated that in all cases, perfect subcarrier and symbol sync has been assumed.

4. Conclusion

In closing, it should be noted that by studying the results given in Figs. 1-9, one can assess typical improvements in telecommunication system performance realizable by using decision-directed feedback in a data-aided carrier tracking loop. Future work will develop design procedures for second-order loops with unequal loop filters.

References

1. Lindsey, W. C., "Design of Block-Coded Communication Systems," *IEEE Trans. Commun. Technol.*, Vol. Com-15, No. 4, pp. 525-534, Aug. 1967.
2. Lindsey, W. C., "Block Coding for Space Communications," *IEEE Trans. Commun. Technol.*, Vol. Com-17, No. 2, pp. 217-225, Apr. 1969.
3. Easterling, M. F., Spear, A. J., and Tausworthe, R. C., *A High-Rate Telemetry System for the Mariner-Mars 1969 Mission*, Technical Report 32-1354. Jet Propulsion Laboratory, Pasadena, Calif., Apr. 1, 1969.
4. Lumb, D., "Test and Preliminary Flight Results on Sequential Decoding of Convolutional Encoded Data from Pioneer IX," *Proceedings of the International Communications Conference*, Boulder, Colo., June 9-11, 1969.
5. Forney, G. D., and Langelier, R., "A High Speed Sequential Decoder for Satellite Communication," *Proceedings of the International Communication Conference*, Boulder, Colo., June 9-11, 1969.
6. Lindsey, W. C., and Simon, M. K., *Data-Aided Carrier Tracking Loops*, Jet Propulsion Laboratory, New Technology Report JPL-19700, NASA-30-11282. Jet Propulsion Laboratory, Pasadena, Calif., Sept. 1969.

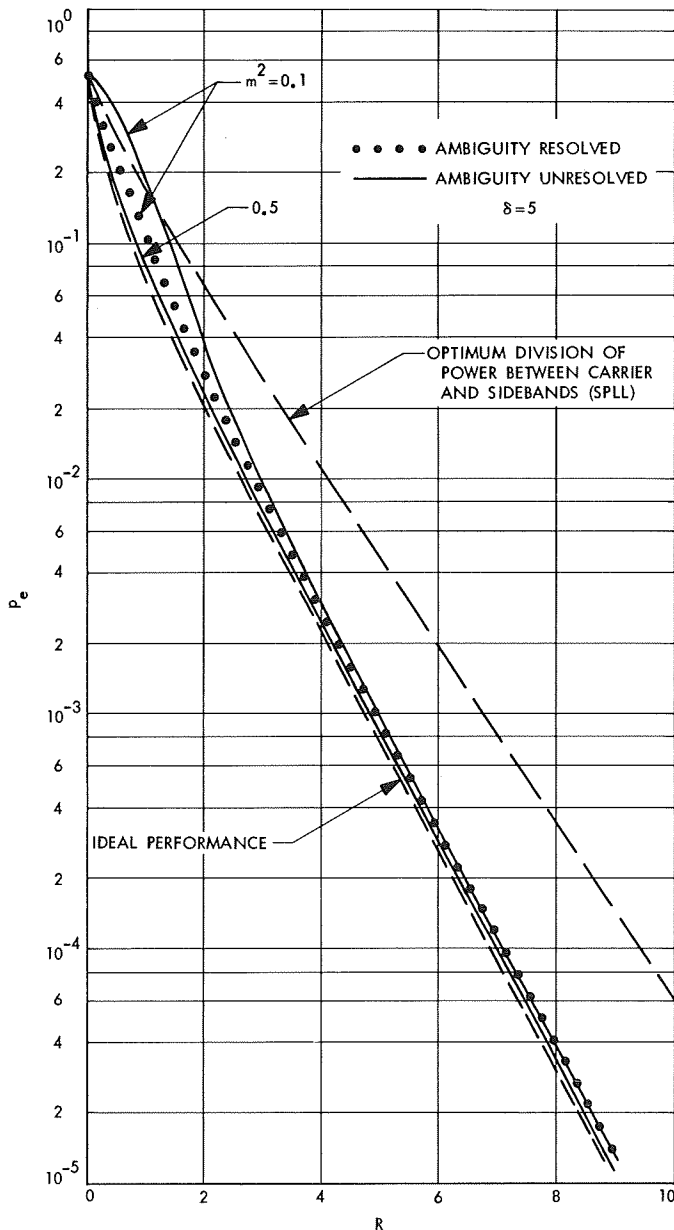


Fig. 9. Comparison of telemetry system error probabilities

7. Lindsey, W. C., "Optical Design of One-Way and Two-Way Coherent Communication Links," *IEEE Trans. Commun. Technol.*, Vol. Com-14, pp. 418-431, Aug. 1966.
8. Lindsey, W. C., "Determination of Modulation Indexes and Design of Two-Channel Coherent Communication Systems," *IEEE Trans. Commun. Technol.*, Vol. Com-15, pp. 229-237, Apr. 1967.
9. Lindsey, W. C., "Nonlinear Analysis of Generalized Tracking Systems," *Proc. IEEE*, Oct. 1969.
10. Davenport, W. B., and Root, W. L., *An Introduction to the Theory of Random Signals and Noise*, McGraw-Hill Book Co., New York, 1958.
11. Viterbi, A. J., *Principles of Coherent Communications*, McGraw-Hill Book Co., New York, 1966.

B. A Comparison of the Performance of Costas or Squaring Loops With Data-Aided Loops,

W. C. Lindsey and M. K. Simon

1. Introduction

In a recent paper,⁵ the authors introduced the concept of data-aided carrier tracking loops. For communications system planning, it is desirable to compare the performance of a data-aided carrier tracking loop, which tracks a suppressed carrier PCM/PSK/PM signal, to the Costas or squaring loops analyzed in Ref. 1 and footnote 6. Consequently, the purpose of this article is to make a rather detailed but brief comparison of the two techniques based upon a criterion of phase error variance. In this article, we draw heavily upon the notation given in Ref. 1 and footnotes 5 and 6.

2. Steady-State Tracking Performance

a. The squaring loop. In Ref. 1 and footnote 6, it was shown that the probability density function (PDF) of the phase-error in a first-order squaring loop is given by

$$p(\Phi) = \frac{\exp[\beta\Phi + \alpha \cos \Phi]}{4\pi^2 \exp(-\pi\beta) |I_{j\beta}(\alpha)|^2} \times \int_{\Phi}^{\Phi+2\pi} \exp[-\beta x - \alpha \cos x] dx, \quad |\Phi| \leq \pi \quad (1)$$

⁵Lindsey, W. C., and Simon, M. K., *Data-Aided Carrier Tracking Loops* (JPL internal document). See also Section A of this chapter.

⁶Lindsey, W. C., and Simon, M. K., *Nonlinear Performance of Suppressed Carrier Tracking Loops in the Presence of Frequency Detuning*, to be published.

where $I_\nu(x)$ is the modified Bessel function of imaginary order and of argument x and

$$\alpha = \frac{\rho}{4} \mathcal{S}_L \quad \beta = \frac{\Omega_0 \alpha}{2B_L} \quad (2)$$

and

B_L = one-sided loop bandwidth

$$\mathcal{S}_L = \left[1 + \frac{2}{SN_0} \int_{-\infty}^{\infty} R_{n_1}^2(\tau) d\tau \right]^{-1}$$

$$R_{n_1}(\tau) = n_1(t) n_1(t + \tau)$$

$$\simeq \frac{N_0}{2} \int_{-\infty}^{\infty} |H_i[j2\pi f]|^2 \exp[j2\pi f\tau] df$$

where $H_i(j\omega)$ is the low-pass equivalent of the input band-pass filter characteristic, $H_i(j\omega)$.

ρ = signal-to-noise ratio in the loop bandwidth of a standard phase-locked loop = $S/(N_0/B_L)$

The parameter \mathcal{S}_L is the so-called loop squaring loss (footnote 6). Herein, all statements made regarding squaring loops are equally applicable to Costas loops because of the stochastic equivalence demonstrated in Ref. 1 and footnote 6.

b. The data-aided loop. For the data-aided loop it has been shown (footnote 5) that the steady-state probability distribution of the phase error is given by

$$p(\phi) = N \exp[-U_0(\phi)] \int_{\phi}^{\phi+2\pi} \exp[U_0(x)] dx, \quad |\phi| \leq \pi \quad (3)$$

where N is the normalization constant and the potential function $U_0(\phi)$ is given by

$$U_0(\phi) = -\frac{2}{K_{00}} \left\{ \Omega_0 \phi + K S^{1/2} \left[\cos \phi \operatorname{erf}(R^{1/2} \cos \phi) + \frac{1}{(\pi R)^{1/2}} \exp(-R \cos^2 \phi) \right] \right\} \quad (4)$$

with

$$R = \frac{ST_b}{N_0}; \quad K_{00} = \frac{N_0 K^2}{2} \quad (5)$$

and

$$\operatorname{erf}(x) = \frac{2}{\pi^{1/2}} \int_0^x \exp(-y^2) dy \quad (6)$$

$$U_o(\phi) = -\frac{4\Omega_o}{N_o K^2} \phi - 2R\delta_e \operatorname{erf} R^{1/2} \left[\cos \phi \operatorname{erf}(R^{1/2} \cos \phi) + \frac{1}{(\pi R)^{1/2}} \exp(-R \cos^2 \phi) \right] \quad (8)$$

It should be clear from footnote 5 that for the special case of suppressed carrier (i.e., $m^2 = 0$), the upper loop gain, K_u , should be chosen equal to zero, while for simplicity of notation we set $K_i = K$. The gain K can be related to the system parameters by

$$K = \frac{2}{\delta_e S^{1/2} T_b \operatorname{erf} R^{1/2}} \quad (7)$$

Figure 10 illustrates the phase error density, $p(\phi)$, for $\Omega_o = 0$, $\delta_e = 5$, and R as a parameter. Notice from this figure that the data-aided device exhibits the same 180-deg phase ambiguity as the squaring or Costas loop.

c. Comparison of the variance of the phase error ($\Omega_o = 0$). If one assumes that an ideal bandpass filter precedes the Costas loop, then (footnote 6)

where $\delta_e = 1/(2B_{Le}T_b)$ and B_{Le} is the effective one-sided bandwidth of the data-aided loop. Substituting Eqs. (6) and (7) into Eq. (4) gives, upon simplification,

$$\sigma_L^2 = \frac{1}{1 + \frac{1}{\rho\gamma}} \quad (9)$$

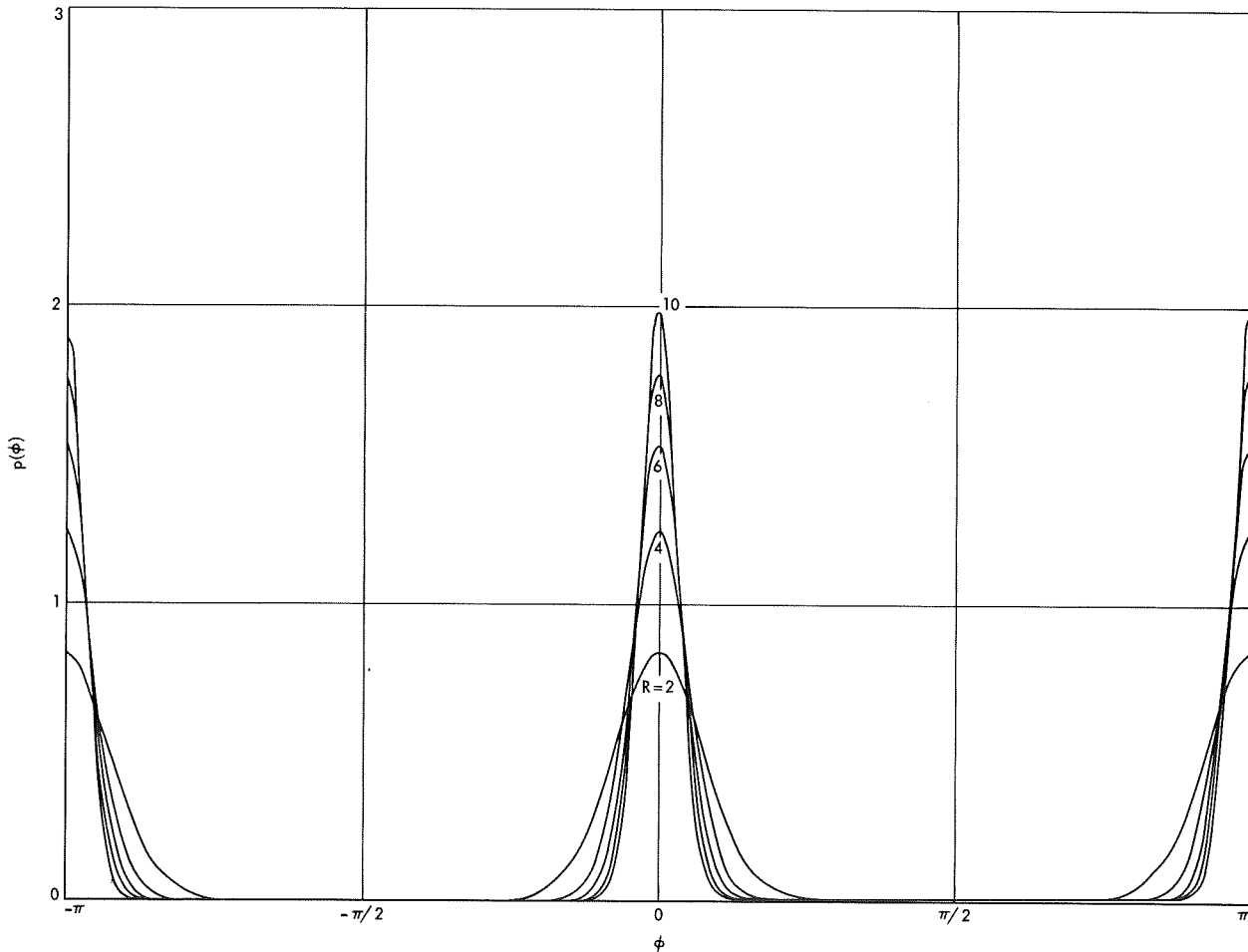


Fig. 10. Phase error density ($\Omega_o = 0$, $\delta_e = 5$)

where $\gamma = B_L/B_i$ and B_i is the one-sided bandwidth of the bandpass filter. Noting that $\rho = 2R\delta$, with $\delta = 1/(2B_L T_b)$, one can easily evaluate (by means of numerical integration) the variance of the phase error using Eqs. (1) and (3). Recalling that $\Phi = 2\phi$ and letting $\delta_e = \delta$; Fig. 11 represents plots of the variance of the phase error versus R for $\delta = 5$ and 20, respectively. The performance of the data-aided loop is independent of γ ; however, the performance of the Costas loop is degraded with decreasing γ . In fact, for all cases examined, the performance of the squaring loop was found to be inferior to that of the data-aided loop. Figure 12 is a plot of the ratio in dB of the variance of the phase error in a squaring loop to that in the data-aided loop for various values of γ with $\delta = 5$. This figure is an alternate representation of the data presented in Fig. 11a.

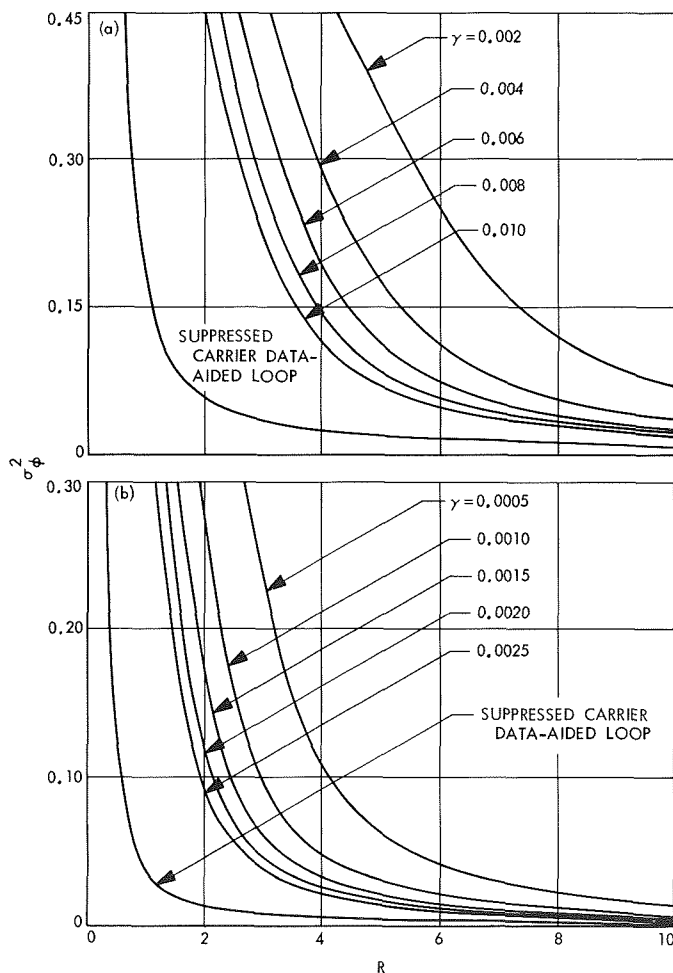


Fig. 11. Variance of phase error versus R assuming an ideal bandpass filter for (a) $\delta = 5$ and (b) $\delta = 20$

Reference

1. Didday, R. L., and Lindsey, W. C., "Subcarrier Tracking Methods and Communication System Design," *IEEE Trans. Commun. Technol.*, Vol. COM-16, No. 4, pp. 541-550, Aug. 1968.

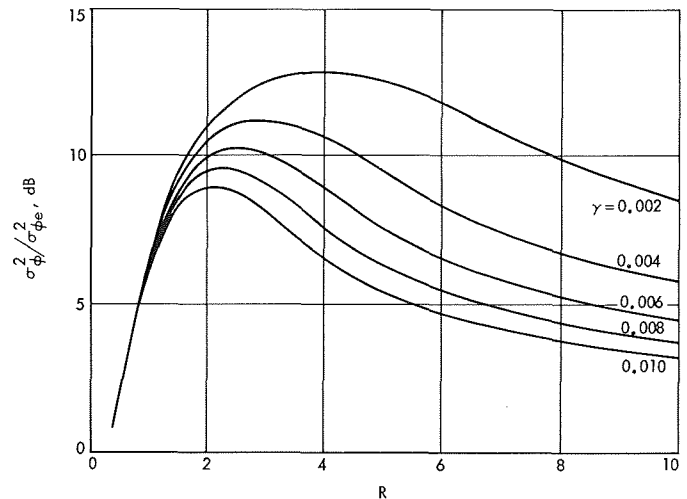


Fig. 12. Ratio of variance of squaring loop phase error to data-aided loop phase error assuming an ideal bandpass filter ($\delta = 5$)

C. Calculations of Power Spectra of Square-Wave Modulated Pseudorandom Sequences, J. H. Yuen

1. Introduction

A Tracking and Data Relay Satellite Network (TDRSN) consisting of relay satellites at high altitude (about 22,000 to 23,000 miles) in synchronous orbit is expected eventually not only to displace a number of ground stations, but also to provide more communications capability than these displaced stations. In one particular area; namely the VHF link that normally serves the smaller, lower orbit, less sophisticated user satellites; the TDRSN concept has placed a significant burden on the design of the TDRS telecommunication system. In addition to the multipath problem that exists between the user and the TDRS, there exists a potentially more severe problem created by the presence of unintentional interference sources located on the earth and found within the VHF bands.

A Pseudonoise (PN) system has been proposed⁷ to overcome the interference and multipath problems associated

⁷Tracking and Data Relay Satellite Network Final Study Report, Sept. 30, 1969 (JPL internal document).

with the VHF communication links between the TDRS and the low data rate user satellites.

In this article, the effect of the phase relationship between the clock and the PN code in this PN system on the capability of the system to reject unintentional interferences is investigated.

The spectral density of the signal $(PN \times \text{clock})$ is shaped by shifting the relative phase between the clock and the PN code such that notch filters(s) may be applied advantageously to reject the unintentional interferences. This is done so that a clear channel can be obtained without sacrificing a significant portion of the total available signal energy to maintain satisfactory system performance.

2. Calculation

The PN generator is driven by a clock at the transmitter of the PN system in the TDRSN. The PN code is multiplied by the clock signal with a phase shift ϕ relative to the PN (Fig. 13). Correspondingly, in the receiver end, the synchronized PN system (without RF) is as shown in Fig. 14.

Let the bit rate of the PN sequence be denoted by T , the clock period by $2T$, and let ϕ be the relative phase

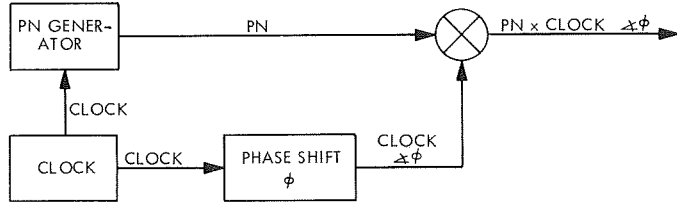


Fig. 13. Transmitter-coder generator

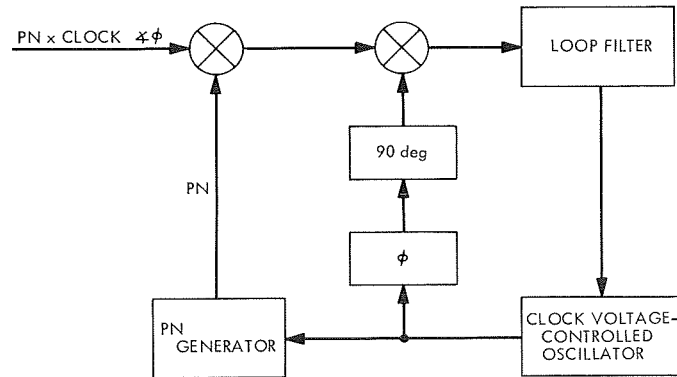


Fig. 14. Receiver-decoder

between the clock and the PN signals (Fig. 15). Then, it may be shown (Ref. 1) that the signal

$$y(t) = [PN \otimes \text{clock} \angle \phi](t) \quad (1)$$

is approximately of a negative equally probable (NEP) process when the period of $y(t)$ is large. The spectrum⁸ of such a process is found (Ref. 1) to be

$$Y(f) = \frac{1}{T} \sum_{i=1}^2 p_i |H_i(f)|^2 \quad (2)$$

where

$$H_i(f) = \int_0^T h_i(t) \exp(-j2\pi ft) dt \quad (3)$$

and $h_i(t)$, $i = 1, 2$, are the basic waveforms in the basic switching interval $[0, T]$.

In the present case, the basic switching waveforms are as shown in Fig. 16. The p_i , $i = 1, 2$, are the probabilities of the waveforms $h_i(t)$, $i = 1, 2$ respectively.

Since,

$$|H_1(f)|^2 = |H_2(f)|^2 = |H(f)|^2 \quad (4)$$

and for a PN sequence, and clock as in Fig. 15,

$$p_1 + p_2 = 1 \quad (5)$$

one can obtain from Eq. (2) the spectral density of $y(t)$,

$$\begin{aligned} Y(f) &= \frac{1}{T} |H(f)|^2 \\ &= \frac{1}{T} \left| \int_0^T -\exp(-j2\pi ft) dt + \int_T^T \exp(-j2\pi ft) dt \right|^2 \\ &= \frac{1}{T\omega^2} (6 - 4\cos\alpha + 2\cos 2\beta - 4\cos(2\beta - \alpha)) \end{aligned} \quad (6)$$

where

$$\left. \begin{aligned} \omega &= 2\pi f \\ \alpha &= \omega\tau = \frac{\omega T\phi}{180} \\ \beta &= \frac{\omega T}{2} \end{aligned} \right\} \quad (7)$$

⁸It should be noted that the spectral densities under consideration are, in reality, line spectrums, but that the line density is so high that their envelopes adequately describe them.

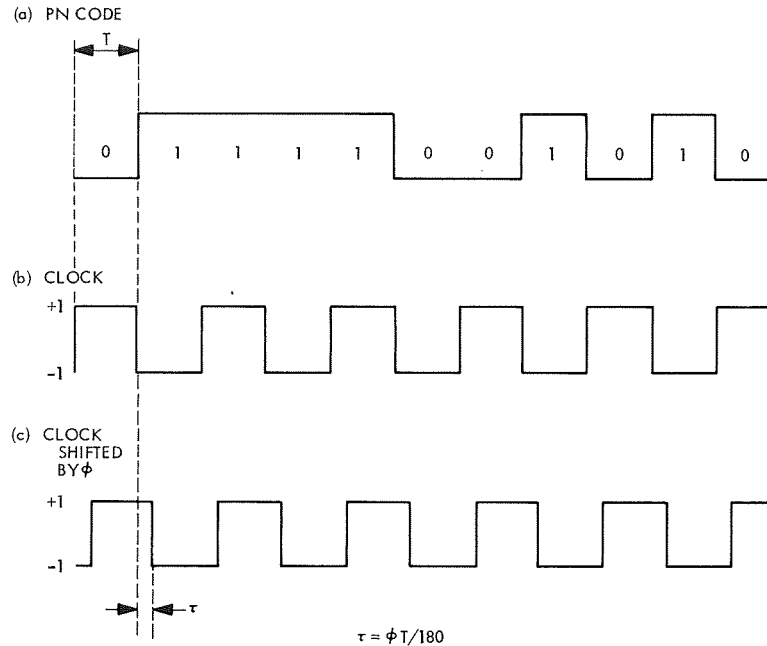


Fig. 15. PN code and clock relationships

The density function $Y(f)$ can be computed for any phase relationship between the PN code and the clock by using Eq. (6). The distribution function (power)

$$F_Y(f) = \int_{-\infty}^f Y(\xi) d\xi \quad (8)$$

can also be calculated. These density and distribution functions (for frequencies within the receiver bandwidth) are plotted in Fig. 17 for various phase shifts $\phi_k = 5k$ deg; $k = 0, 1, 2, \dots, 36$. In these calculations, parameters values of $T = 1.25 \times 10^{-6}$ s and receiver bandwidth = 1.6 MHz as indicated in footnote 7 have been used.

3. Conclusion

The density spectrum in Fig. 17a indicated that when $\phi = 0$ and 90 deg, notch filtering technique might be applied advantageously to reject the unintentional interferences. However, when $\phi = 90$ deg, only about 72% of the total power within the 1.6 MHz bandwidth is preserved as compared to that when $\phi = 0$ deg, i.e., PN and clock are in phase as shown in Fig. 17b. Such a power loss

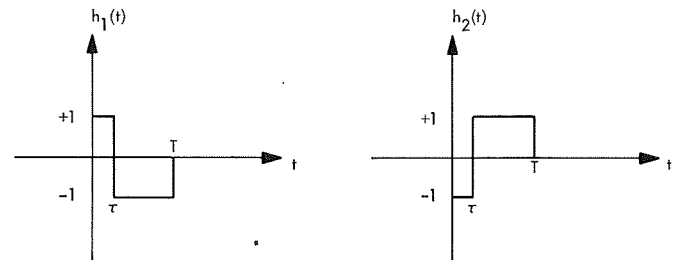


Fig. 16. Basic switching waveforms in one code bit T

would significantly degrade the system performance. Other phase relationships do not offer any advantage over that when $\phi = 0$ deg; therefore, when the clock and the PN signal are in phase, the PN system offers the best anti-interference and anti-multipath capability under the present system and channel constraints.

Reference

1. Tittsworth, R. C., and Welch, L. R., *Power Spectra of Signals Modulated by Random and Pseudorandom Sequences*, Technical Report 32-140, Jet Propulsion Laboratory, Pasadena, Calif., Oct. 10, 1961.

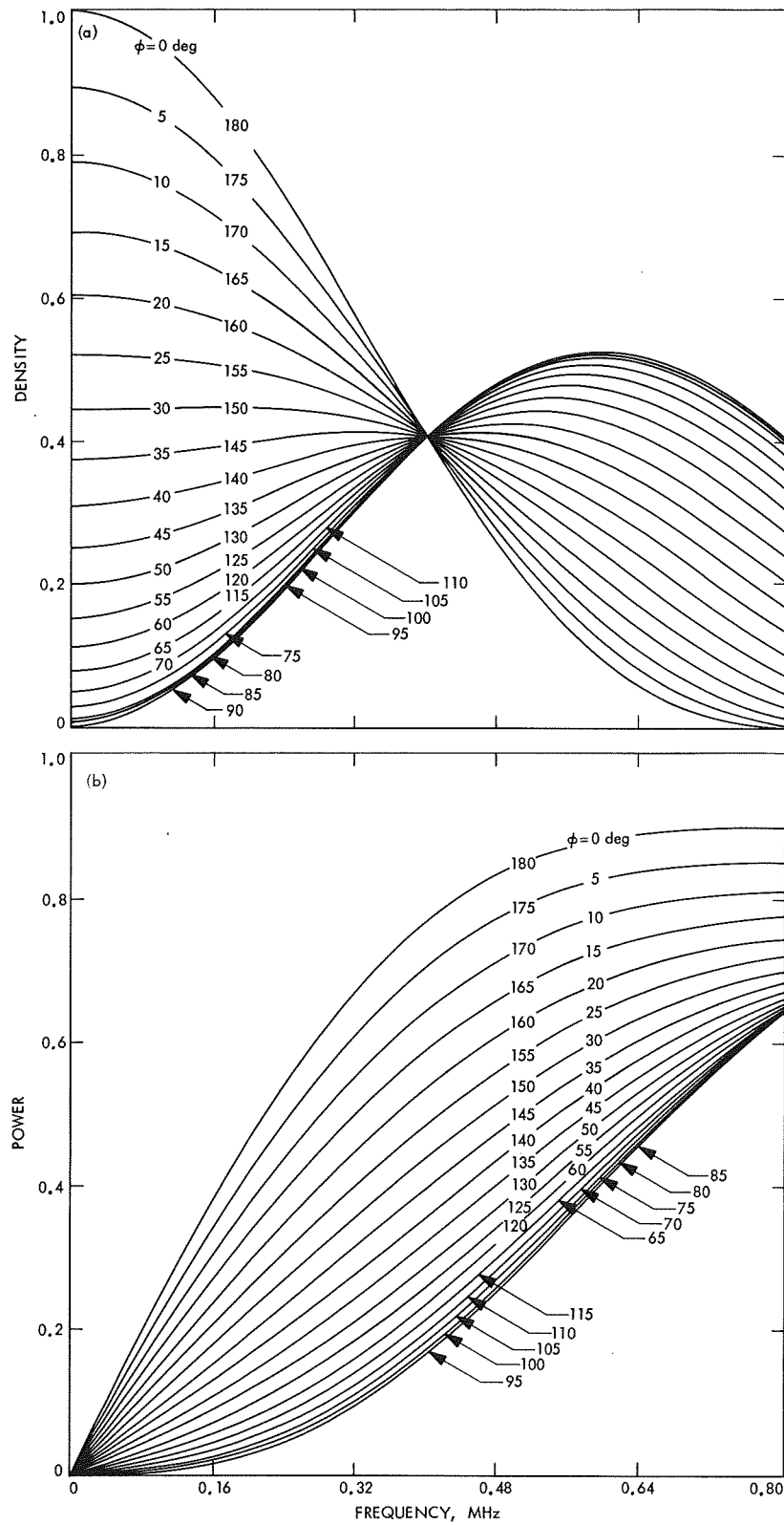


Fig. 17. PN code and clock density and distribution functions:
(a) spectral density, and (b) cumulative power spectrum

IX. Spacecraft Power

GUIDANCE AND CONTROL DIVISION

A. Photovoltaic Radiation Effects, *P. A. Berman*

1. Introduction

One of the major concerns in the use of silicon solar cells for space is the fact that significant decreases in power can occur as a result of exposure to protons and electrons. It has been found that silicon cells with lithium doping can exhibit considerable annealing of radiation damage at temperatures of 28°C or greater. A comprehensive program has been established at JPL to study the properties of lithium-doped silicon and to investigate the applicability of lithium and other dopants to the radiation hardening of silicon solar cells. The results of some of the JPL-funded work in this area are discussed below.

2. Program Results

RCA has irradiated a total of 32 lithium-doped cells fabricated from crucible-grown material to a fluence of 1×10^{14} 1-MeV electrons/cm². These cells have exhibited no redegradation of recovered power 69 days after irradiation. All cells, except those from lot C2, exhibited significant recovery. In some cases, the recovered efficiencies were higher than those of N/P (state-of-the-art) cells

similarly irradiated. The cells from the C2 lot are quite interesting in that the silicon ingot from which they were fabricated was doped with antimony rather than phosphorus (as were all the other cells thus far investigated).

Measurements made by RCA on an experimental lot of cells (fabricated from float-zone material), having undergone various lithium diffusion and redistribution schedules (lot C4), indicated that an increase in minority carrier diffusion length and a decrease in lithium density occurs with increasing redistribution time. However, the longest minority carrier diffusion length of any of the subgroups (i.e., having a given lithium diffusion schedule) was only 40 μ m, which is still relatively short. Cells within each subgroup were similar in behavior. The cells with the best initial performance suffered the greatest fractional power degradation during irradiation; however, all groups recovered to within 4% of the original power after irradiation to 10^{14} 1-MeV electrons/cm².

A type of redegradation of cells fabricated from float-zone material has been observed by RCA in which open-circuit voltage and curve shape degrade, while short-circuit current and diffusion length remain constant. Thus,

this type of degradation appears to be totally unrelated to minority carrier diffusion length degradation. After exposure to a fluence of 3×10^{15} 1-MeV electrons/cm², the relative open-circuit voltage degradation after 11 days annealing at room temperature was greater than the remaining damage to short-circuit current. There is a strong indication that poor junction properties and contacts might be more common causes for problems in lithium cells than changes (degradation) in minority carrier lifetime.

TRW has exposed groups of lithium-doped float-zone silicon solar cells to fluences of 3×10^{14} and 3×10^{15} 1-MeV electrons/cm² and subsequently stored them at room temperature. Half recovery times of 0.5 h and 2 to 3 h were observed for the respective fluences. Maximum annealed values were higher for the lower fluences. Experiments on lithium-doped crucible cells indicated that although the annealing times were considerably longer, both the initial outputs and the maximum annealed outputs were higher than the float-zone cells investigated. It has been found that storage of lithium-doped cells fabricated from crucible-grown silicon at temperatures between 60 and 100°C produces annealing times comparable to those of moderately lithium-doped float-zone cells stored at 28°C.

Investigations have been carried out by TRW on silicon diffused with lithium by means of a lithium aluminum hydride source. It was found that lithium-diffused silicon having concentrations between 10^{15} and 10^{18} lithium atoms/cm³ exhibited Hall coefficients similar to those previously obtained for arsenic-doped silicon and silicon doped with lithium in the melt. Carrier removal studies of the lithium-diffused silicon also indicated an exponential dependence of the carrier removal rate on the carrier concentration, as was observed previously in lithium melt-doped silicon.

It has been found by TRW that non-lithium-doped P/N solar cells exhibit an 0.17-eV activation energy after irradiation, which is believed to be the A-center (oxygen-vacancy pair). Float-zone and crucible-grown lithium-doped solar cells were found by TRW to exhibit a pre- and post-irradiation activation energy of approximately 0.06 eV. The absence of the 0.17-eV A-center activation energy immediately after irradiation and prior to annealing indicates that the simple model of A-center annihilation through lithium diffusion might be inadequate to explain the kinetics of radiation annealing in lithium-doped silicon. It is postulated by TRW that lithium might be involved in the initial defect site in some manner such as lithium-oxygen-vacancy or a lithium-vacancy complex.

Northrop Corporate Laboratories has been studying the effects of various dopants in silicon on the resultant radiation characteristics, primarily with respect to minority carrier lifetime. The first series of irradiations was performed in a Febetron 705 electron-beam machine with an average electron energy of 1.7 MeV. Damage coefficient determination by means of minority carrier lifetime measurements on irradiated samples indicated that aluminum-doped float-zone silicon fabricated by General Electric was a factor of four more radiation-resistant than boron-doped crucible-grown silicon normally used in solar cell fabrication. The results were somewhat suspect, however, because the damage coefficient of phosphorus-doped crucible-grown silicon appeared to be almost identical to that of the boron-doped silicon, in contradiction to results obtained by other workers which show the boron-doped material to be less radiation sensitive. Further investigation of the experiment indicated that the Febetron has an extremely high component of low-energy electrons and that the lifetime measurements were strongly influenced by surface damage effects on the samples, despite efforts to avoid this through design of sample configuration. Subsequent exposure of additional samples to Co⁶⁰ gamma irradiation, which yields relatively uniform damage throughout the samples, indicated little, if any, damage coefficient superiority of the aluminum-doped samples over the boron-doped samples for this type of radiation (i.e., radiation which yields point defects uniformly throughout the sample).

B. Photovoltaic Supporting Development, R. K. Yasui

1. Introduction

The development of photovoltaic devices has progressed over the past decade and such figures of merit as watts per pound and dollars per watt for individual solar cells have been significantly improved. Unfortunately, these improvements, although significant when applied to the individual cell weight and costs, do not have the same impact on the complete solar cell array. The present-day array designs and construction techniques result in an array weight about 10 times the weight of the essential active device, and results in costs that are increased by more than a factor of 2. Obviously, if progress is to be made in the reduction of weight and cost of silicon arrays, it must include studies of all components in the array and not be restricted toward the solar cell device alone. As a result, new materials and techniques are being investigated by JPL under two separate contracts to determine the proper approach for making a significant improvement in array design. New array concepts, along with

unique fabrication methods, are being studied, the goals being increased power-to-weight ratios and decrease in the cost per watt of power generated.

2. Integrated Lightweight Flexible Solar Array

A program was initiated by JPL in June 1969 under Contract 952560 entitled, "Development of an Integrated Lightweight Flexible Silicon Solar Cell Array," with Heliotek, a Division of Textron, Inc. The purpose of this program is to analyze, develop, and test the various components and material systems which can be integrated to form a lightweight flexible solar cell array capable of producing 100 W/lb at a price of \$100 per watt or less. To achieve the objectives of this program, the array study has been divided into seven separate tasks, five of which are directly concerned with the development of the solar cell array. The remaining tasks are concerned with demonstration modules and program documentation. A literature search and review was performed by Heliotek to determine the current state-of-the-art development of lightweight arrays to minimize duplication of effort and to clarify the approach to be taken.

To obtain useful cell price information, a survey was made to determine the cost as a function of size, thickness, quantity, and configuration. An analysis was conducted based upon prices from 1958 through 1969 along with some current data. A learning-curve-type trend became apparent and, by utilizing this information, estimated cell prices for the present and a future (1974) period were generated. This allowed a comprehensive determination to be made of power, weight, and cost versus thickness, and an optimum cost effective cell design was identified. The optimum cell presently available was found to be 8 mils thick. By 1974, it is estimated that 4-mil solar cells will become optimum.

A preliminary analysis of a continuous glass ribbon cell coverslide has been completed. The analysis was divided into two parts, adhesive selection and coverglass selection and sizing. Several adhesives were evaluated, and Sylgard 182 or Dow Corning XR6-3489 was determined to be the most satisfactory adhesive. A thin Corning 8871 high-lead-content ribbon glass was evaluated and methods of applying the glass to the cell were developed. It was determined that the ribbon configuration could be supplied in various chemical compositions, including 0211 microsheet which is commonly used for silicon solar cell covers, thus minimizing the need for a comprehensive radiation evaluation program. The problems associated

with handling and cutting glass were difficult and required special techniques for their solution. The approach was to bond the thin glass to the solar cell and trim the glass to the proper dimension after the bonding adhesive had cured. Three methods of trimming the glass were evaluated. They were a hot nichrome wire, a small precision sandblasting unit, and a diamond scribe. This evaluation resulted in the diamond scribe being the most acceptable technique and the hot wire the least acceptable.

An analysis of solar cell bus bar interconnector design parameters is also being studied. This analysis will consider stress distributions, thermal cycling fatigue, and vibration fatigue. Preliminary results of the analysis indicate that for lightweight flexible arrays the stress loop type interconnector was found to be a marginal design for fatigue resistance. Consequently, alternate interconnector concepts are being investigated. A wrap-around type utilizing thin film technology is being considered and appears desirable for the lightweight flexible array.

An analysis of the lightweight ribbon substrate showed that the bond strength was essentially independent of bond area or ribbon width. This allows one to design a substrate using narrow ribbons rather than completely covering the back of cells, thus saving considerable weight. This system appears to be attractive in that it not only eliminates the unnecessary substrate material, but also eliminates much of the cell bonding adhesive that does not contribute to the structural integrity of the solar cell array. This analysis was begun by investigating the failure mechanism of the ribbon substrate concept. The results of the analysis indicated that the tensile strength of cells bonded onto a polyimide ribbon with RTV silicone rubber increased to a point, then became relatively independent of the distance that the polyimide ribbon overlapped the cell. Optimum overlap was found to be about 0.10 in.

These new ideas for materials and methods were then applied to the overall assembly to improve the design in respect to both superior performance and more economical manufacturing possibilities. The net result anticipated from this development study will be an improved overall solar cell array, which will be versatile and will yield useful engineering data that can be applied to all solar power systems. The ultimate objective in this study will be to develop an advanced array made up of an integrated module of silicon solar cells about one foot square or larger which can be readily interconnected at low cost to make up a larger system.

3. Electroformed Aluminum Silicon Solar Cell Contacts

A program was initiated by JPL in June 1969 under Contract 952485 entitled, "Research and Development of Electroformed Aluminum Solar Cell Contacts and Interconnects," with Electro Optical Systems, Inc., a Division of Xerox Company. The program is divided into three main tasks: (1) development of electroformed aluminum contacts, (2) development of electroformed aluminum intercell connectors, and (3) fabrication of multicell matrix. These tasks are to be accomplished consecutively. Most of the effort in this reporting period was devoted to task 1. The primary program objective is to develop a one-step process of electroforming the contacts and interconnections onto state-of-the-art N/P silicon solar cell wafers. The achievement of such a goal will reduce solar array costs. The resultant solar array will be lighter in weight and will be capable of withstanding higher and lower temperature extremes, since no soldering will be required. The technical goals of this program are to improve solar cell performance and increase reliability of solar cell contacts and interconnects. The selected solar cell contact configuration is shown in Fig. 1.

All process parameters for fabricating the electroformed aluminum-contact solar cells from N/P silicon wafers were individually investigated. The individual parameters, including masking, electrodeposition, and silicon monoxide antireflection (AR) coating, have been optimized and integrated into producing desirable electroformed aluminum-contact solar cells with exceptionally good yield. The results of these studies led to the following conclusions:

- (1) The photoresist technique produces the most durable and definitive cell contact mask.
- (2) The most desirable aluminum contact material is obtained when it is electrodeposited under the following conditions:

Aluminum bath composition:

AlCl_3 : 3.4 molar

LiAlH_4 : 0.3 molar

Solvent: anhydrous ether

Current density: 15 A/ft²

Temperature: 22 to 25°C

- (3) A contact adhesive strength of >2,000 g (Hunter pull tester) is obtained when the aluminum electrodeposit on the silicon wafer is heated for 2 min in a 350°C oven.

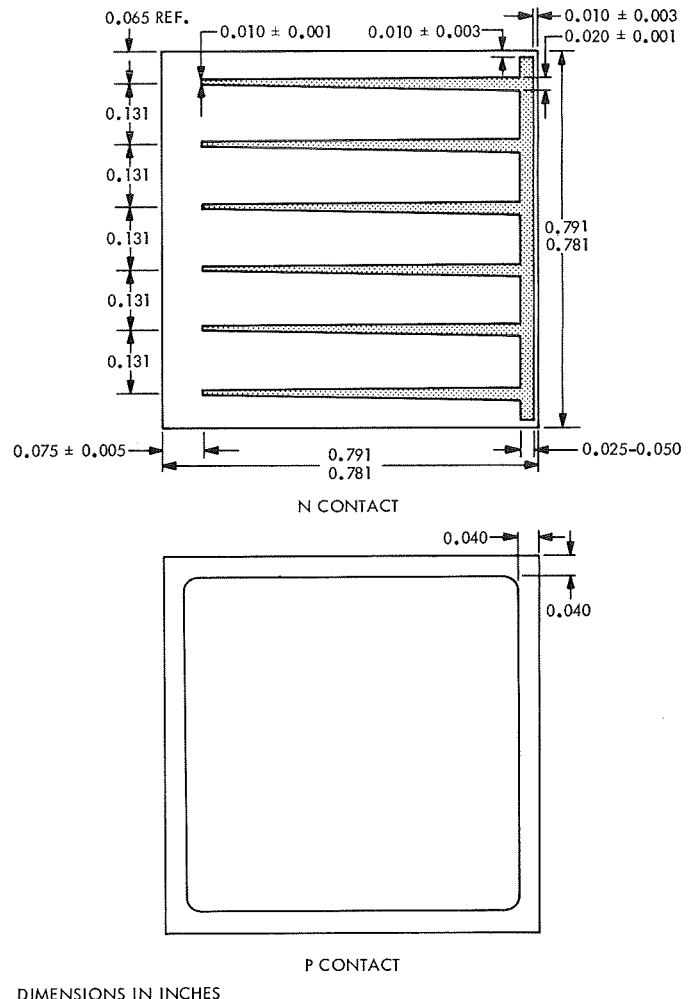


Fig. 1. Aluminum electroformed silicon solar cell with tapered grid lines

- (4) The cell peak power increases 15 to 17% (η increases $\approx 1.8\%$) when a 1400-Å layer of silicon monoxide AR coating is applied to the bare electroformed aluminum-contact solar cells.
- (5) The "weak knee" and open-circuit voltage of the electroformed aluminum-contact solar cells can be improved by etching the edges of the cell in hydrochloric acid, which removes the minute amount of aluminum electrodeposited across the N/P junction.

Electroformed aluminum-contact solar cells have resulted in efficiencies of greater than 10% measured under simulated air mass, zero solar spectrum, and 28°C temperature. A typical current-voltage characteristic curve of an aluminum electroformed contact solar cell is shown in Fig. 2.

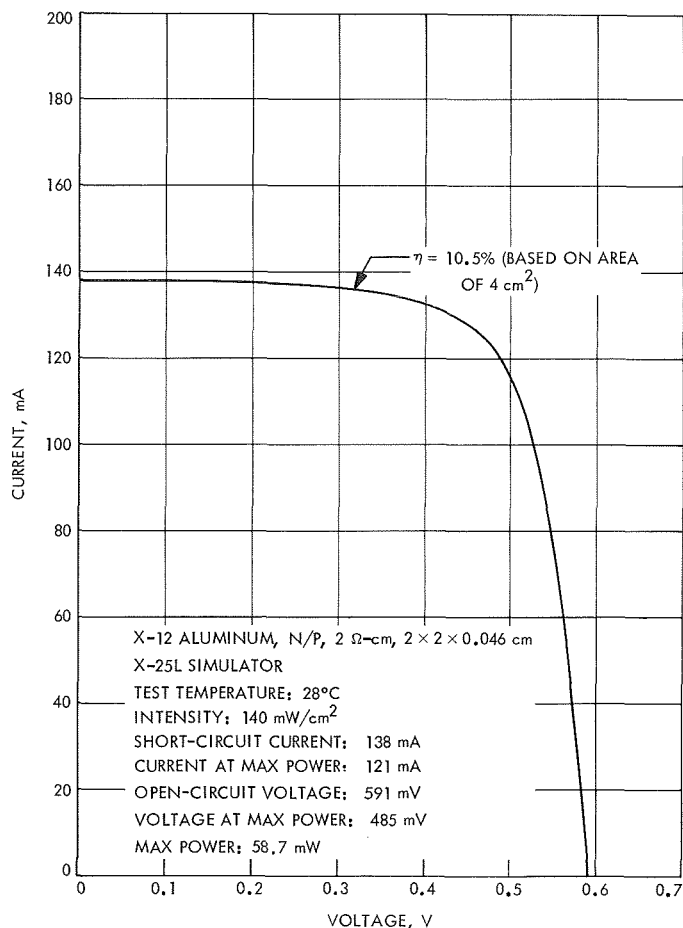


Fig. 2. Aluminum electroformed contact solar cell current-voltage characteristic curve

C. A Circuit Design for Eliminating High Current Spikes at Converter and Inverter Inputs, C. McLyman

1. Problem Statement

An often-used inverter circuit for obtaining an ac voltage from a dc source is shown in Fig. 3. The circuit consists of a low-power saturating oscillator driving a power amplifier.¹ The primary advantage of the circuit is that the saturating transformer in the oscillator can be made quite small since only a few watts are required to drive the power amplifier. This results in high inverter efficiency.

¹In Fig. 3, C1R1 and C2R2 form speedup networks to improve switching times. R5 and R4 are equalizing resistors which prevent excessive drive. C3 and L1 are input line filters. This dc-to-ac inverter can be converted to a dc-to-dc converter by simply adding rectifying diodes and an appropriate filter on the transformer output.

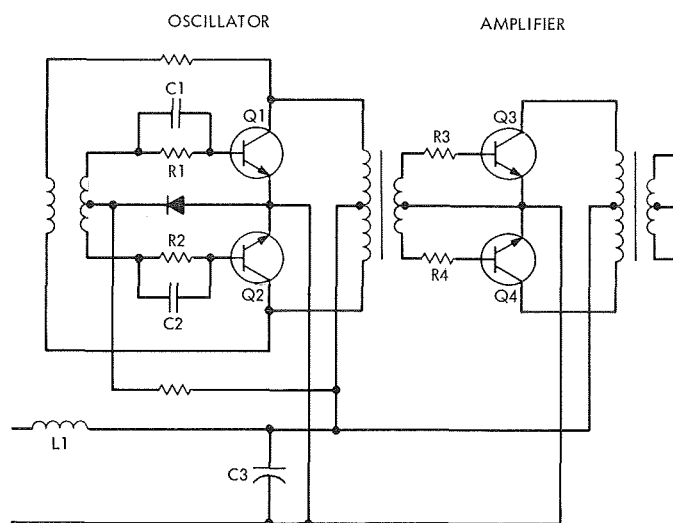


Fig. 3. Dc-to-dc converter with standard drive circuit

ciency. Also, since the oscillator transistors switch low current, the dangerous region of secondary breakdown can be avoided.

This type of oscillator circuit, although quite efficient, has one major disadvantage. It tends to overstress power amplifier transistors Q3 and Q4 and to make the line filter L1 and C3 quite large. This stress on the power amplifier transistors and increased line filter size results from the current spikes generated by having both switches (Q3 and Q4) turned on at the same time for brief periods during the switching cycle.

This simultaneous turn-on of the power transistors can be explained as follows: The turn-on and turn-off of transistors Q3 and Q4 are comprised of four time intervals: t_d = delay time, t_r = rise time, t_s = storage time, and t_f = fall time. The turn-on time is the sum of t_d and t_r , the turn-off time is the sum of t_s and t_f . The turn-off to turn-on ratio of the power switching transistors is approximately 3. This means the transistor turns off 3 times slower than it turns on and during a small part of each switching cycle both transistors will be conducting. When both transistors are conducting, the power source to the inverter is shorted and the input current is limited only by the beta of the transistors. When the transistors cease to conduct simultaneously, i.e., one transistor finally turns off, a large voltage spike appears from the collector to emitter of the turned-off transistor due to the interruption of the high current flowing through the transformer leakage reactance.

2. Problem Solution

The large current surges and associated voltage spikes could be avoided if the turn-on of the nonconducting transistor could be delayed until the conducting transistor is completely turned off. Thus, simultaneous conduction of both transistors would not occur.

To accomplish this, a "volt-second reactor" was developed for insertion into the base drive circuits of the power amplifier (Fig. 4). This device absorbs the leading edge of the square-wave drive from the oscillator for an amount of time equal to $(t_d + t_r) + (t_s + t_f)$. In Fig. 4, R3 and R4 limit the drive current to the transistor switches and R5 and R6 supply the magnetization current to the "volt-second" device. The voltage drop across R5 and R6 due to the magnetization current has to be less than the transistor V_{BE} turn-on value. The cores are reset with a dc bias winding.

In designing the "volt-second" device and its associated circuitry, the following design equations are useful. The number of turns N can be calculated from the usual magnetic expression:

$$N = \frac{\Delta e \Delta t}{A 2B_m \times 10^{-8}}$$

where Δe = voltage, N = turns, A = core area (cm^2), B_m = flux density, and Δt = time (s). The magnetization current I_m is found from

$$I_m = \frac{H_{m1}}{0.4\pi N}$$

where H_{m1} = magnetizing force. Finally, the maximum value of the resistors R5 and R6 must not exceed the value from the expression

$$R_5 = R_6 = \frac{V_{BE(\min)}}{I_m}$$

where $V_{BE(\min)}$ = minimum voltage required to turn on the transistors Q3 and Q4.

3. Experimental Results

To test the design theory and equations, a 30-W dc-to-dc converter was designed and breadboarded. The breadboard was such that the "volt-second" device could be easily inserted or removed for test purposes. The input and output voltages of the converter were 30 Vdc.

Figures 5 and 6 show the oscillograms of the input current surges to the converter and the collector-to-emitter

voltage waveforms of one of the switching transistors prior to inserting the "volt-second" device. As can be seen, excess current spikes of up to 2 A were observed and voltage spikes across the transistors reached 40 V. After inserting the "volt-second" device, both the current and voltage spikes were reduced to insignificant values, as can be seen in Figs. 7 and 8.

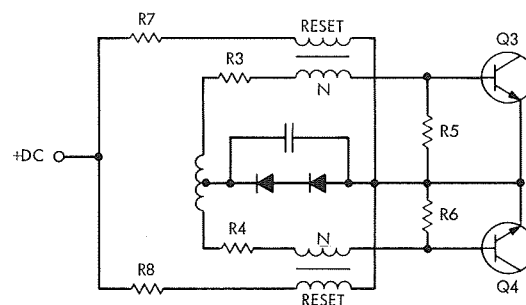


Fig. 4. Dc-to-dc drive circuit with "volt-second" device added

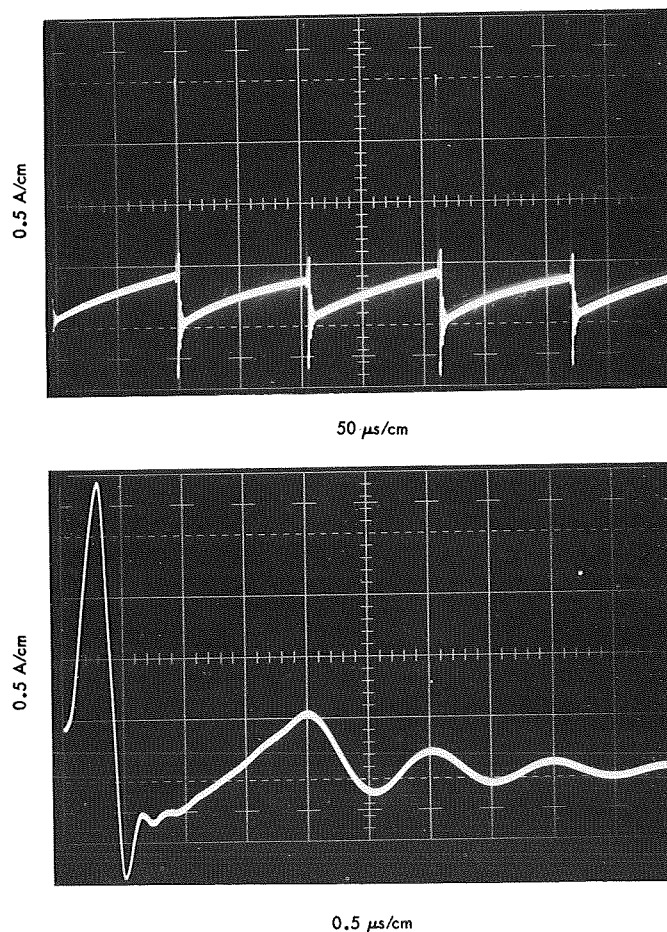


Fig. 5. Typical waveform of input current spikes

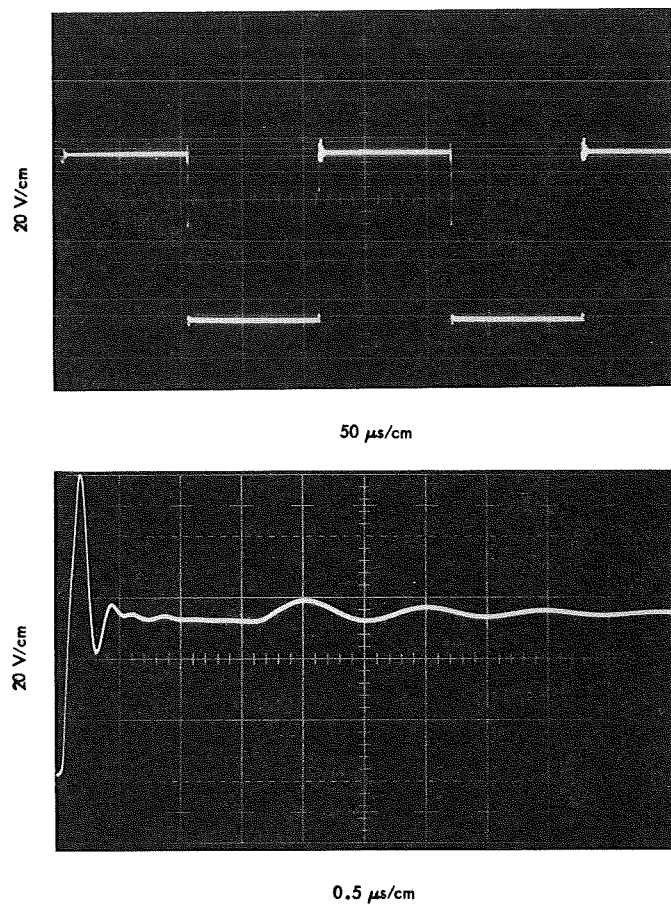


Fig. 6. Typical waveform of collector-to-emitter voltage

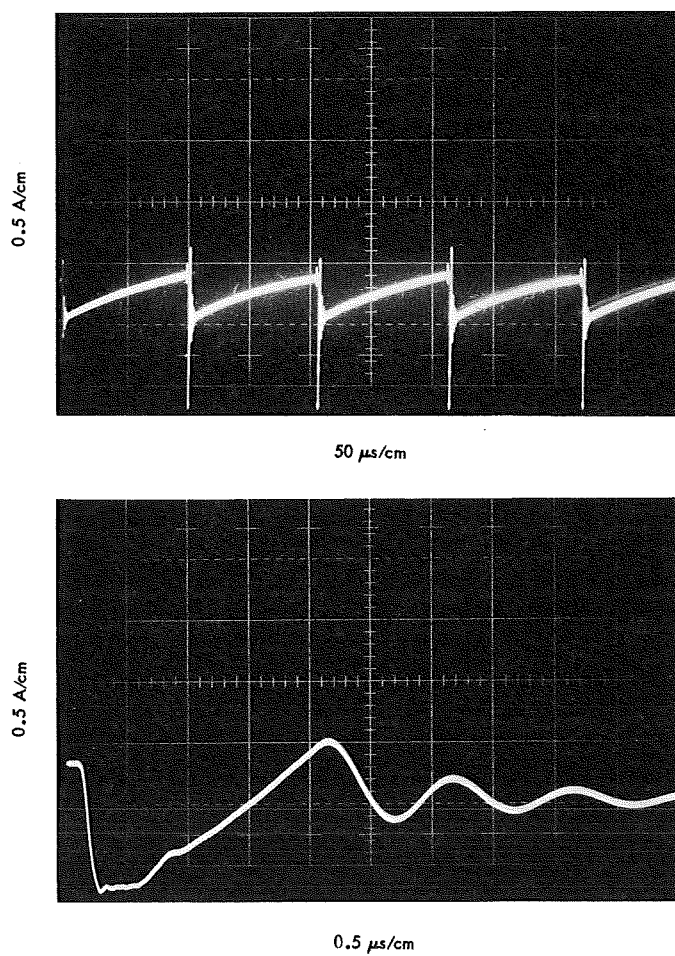


Fig. 7. Current waveform after inserting "volt-second" device

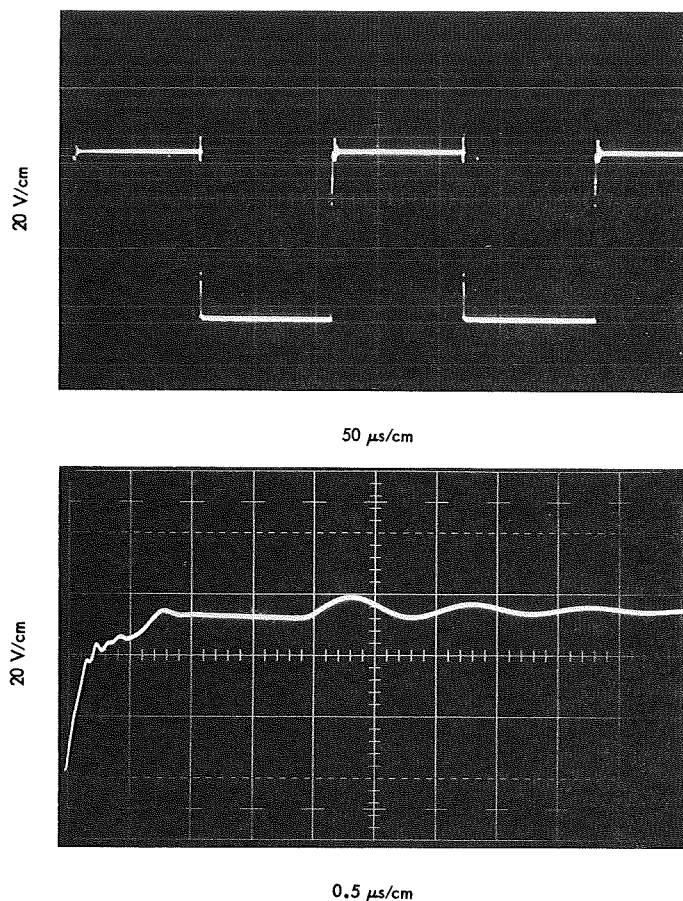


Fig. 8. Voltage waveform after inserting "volt-second" device

Although it is difficult to relate these reductions in current and voltage transients to specific increases in converter or inverter lifetimes, it seems reasonable to assume that their elimination can only aid in improving component reliability.

D. Quad Redundant Shunt Regulator for Nuclear Power Subsystem, D. Hopper and H. Wick

1. Introduction

This article describes the design of a shunt regulator developed for a nuclear power subsystem. This is a result of the activities under JPL contract 952536 to design, develop, breadboard, and test a radioisotope thermoelectric generator (RTG) power subsystem capable of supporting missions to the outer planets.

2. Description

The primary function of the shunt regulator is to regulate the spacecraft primary bus voltage in such a way

that the electrical power drawn from the RTG power source is constant regardless of spacecraft operating mode. The operation of the shunt regulator at any instant of time depends on the difference between the instantaneous power available from the RTG and the spacecraft power demand. This power difference is dissipated within the shunt regulator.

The electrical performance characteristics of a typical Si-Ge power source (Fig. 9) show that optimum loading of an RTG, to minimize hot junction temperature excursions and thus promote long life of the RTG, could be obtained by operating at the maximum power point at all times. This approach would require a maximum power point tracker shunt regulator. Unfortunately, maximum power point trackers are fairly complicated devices. As can be seen from Fig. 9, if a slight relaxation is made on the requirement for a constant hot junction temperature, a constant voltage shunt regulator can be employed to load the RTG and at the same time regulate the spacecraft primary voltage. This is the approach that was taken in

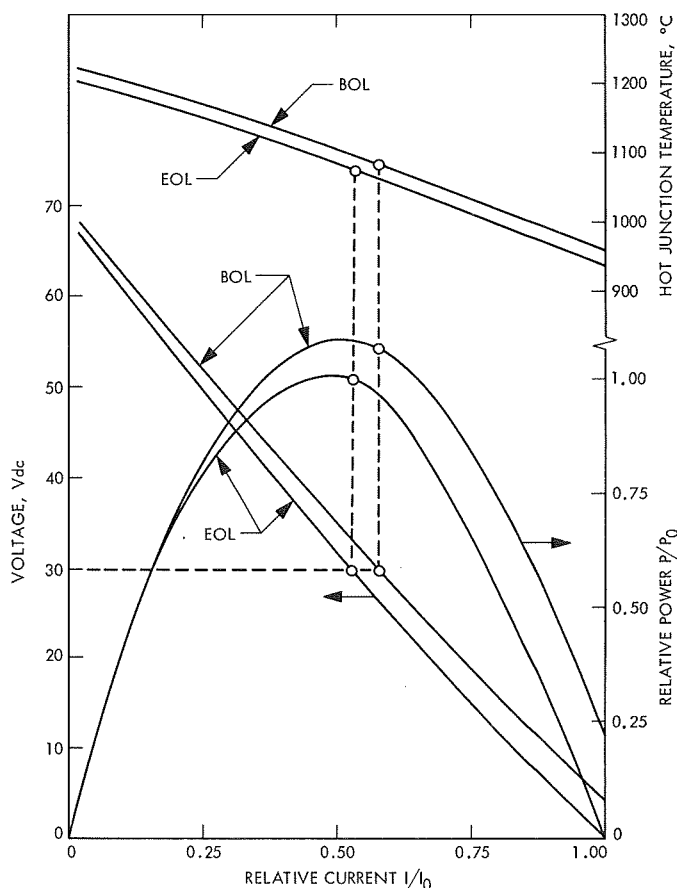


Fig. 9. Typical RTG performance characteristics

the present shunt regulator design. Also note from Fig. 9 that if constant voltage loading is used, the hot junction temperature operating point varies over a small range from beginning of life (BOL) to end of life (EOL).

The design requirements for the shunt regulator of the nuclear power subsystem are listed in Table 1.

The schematic diagram for the quad redundant shunt regulator of the nuclear power subsystem is given in Fig. 10. The design consists of a series-parallel bank of power resistors (R1-R18) in series with quad-H configured power regulators. The control portion of each power regulator is an integrated circuit μ A723 which includes a voltage reference and differential comparator. The output of the μ A723 (A1) is level-shifted by a zener diode (CR1) and then current-amplified by a power transistor (Q4). The power transistors (Q1, Q2, and Q3) conduct current in response to the error signal from A1 to maintain regulation of the primary bus voltage. Each of the four power regulators are connected such that the shunt regulators can tolerate multiple component part failures and still continue to function normally.

3. Test Results

The prototype design shunt regulator has been bread-boarded and tested under simulated flight conditions. It has been observed that all electrical requirements of the shunt regulator as specified are satisfied by the present design. Test data showing regulation as a function of temperature and load current have been plotted and are

Table 1. Shunt regulator design requirements

Parameter	Requirement
Maximum RTG open-circuit voltage, Vdc	70
Maximum shunt current, A	10
Output voltage, Vdc	30
Regulation (includes initial adjustment tolerance, input line variation, load range, temperature, long-term drift, and radiation), %	± 1
Efficiency, %	98 at EOL
Reliability	> 0.9 for 12 yr
Dynamic impedance, Ω	0.1
Transient response, kHz	10
Maximum shunt power, W	320
Temperature range, $^{\circ}\text{C}$	-20 to $+80$
Partial survival capability Shunt regulator shall perform satisfactorily under all conditions, including multiple component part failures within the shunt regulator.	

shown in Fig. 11. Voltage regulation is about $\pm 0.1\%$. Figure 12 gives the results of extensive output impedance testing of the shunt regulator. Load transient response for the present design is approximately 1 V/ms. The computed reliability of the shunt regulator is 0.94 for a 12-yr mission life.

Fig. 10. Quad redundant shunt regulator 

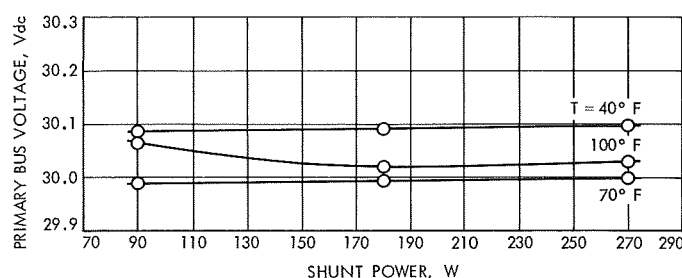


Fig. 11. Voltage regulation characteristics for quad redundant shunt regulator

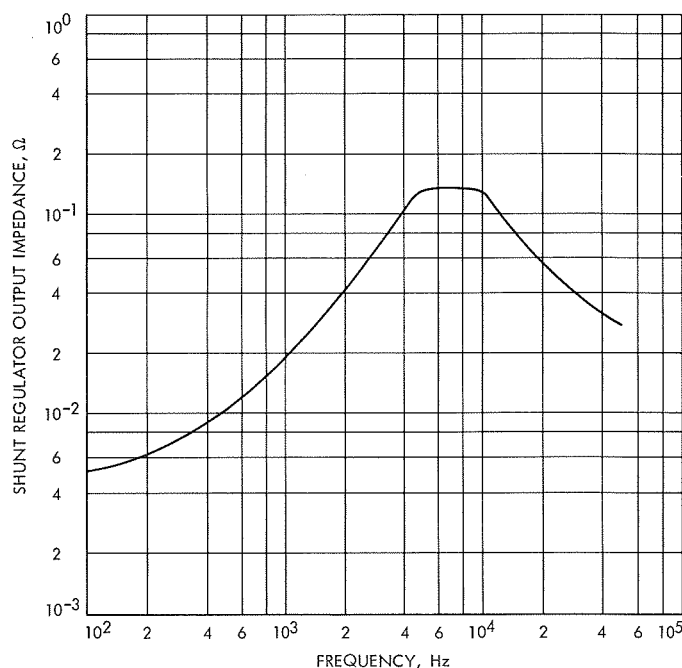
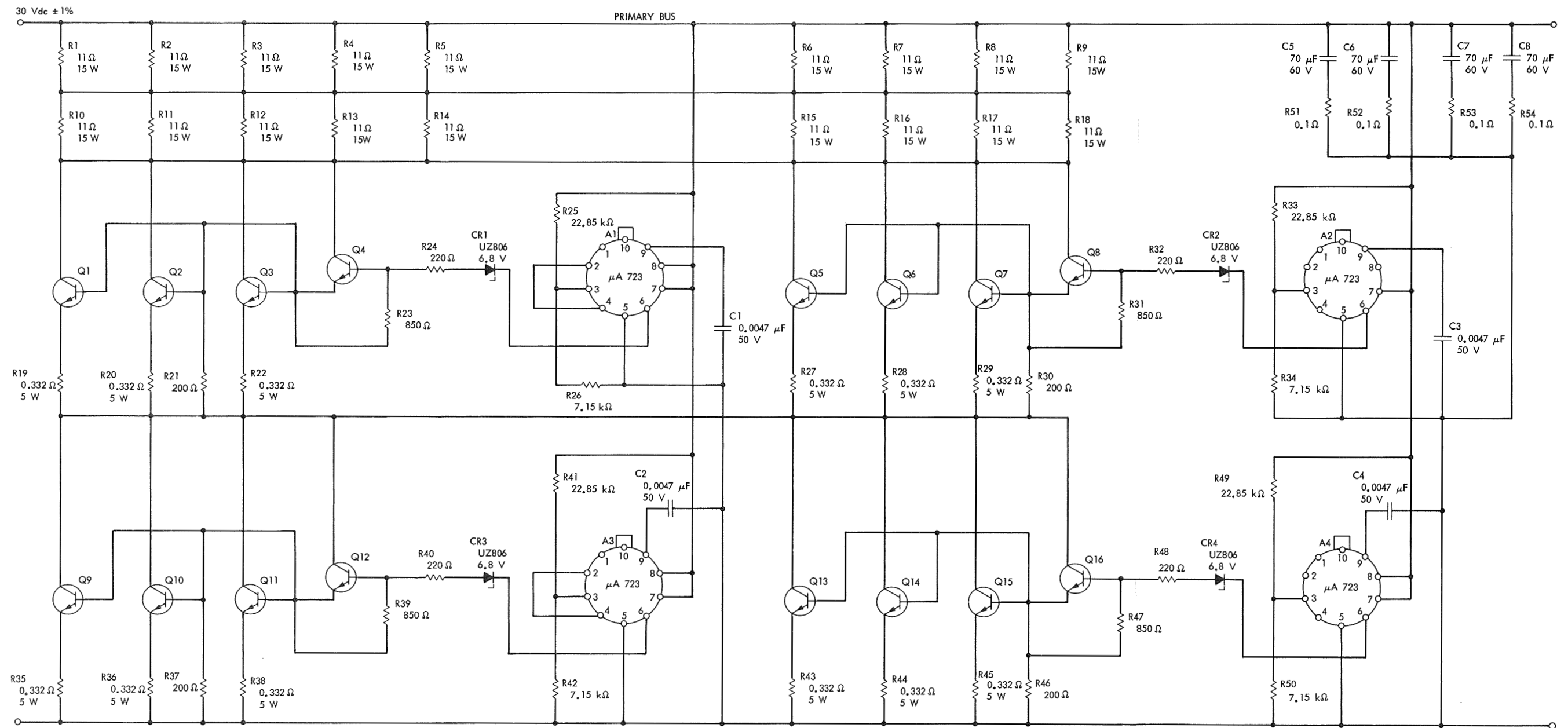


Fig. 12. Output impedance characteristics



Page intentionally left blank

E. Battery Storage Optimization and Design

Studies, R. S. Bogner and R. E. Patterson

1. Introduction

Although a great deal of data has been generated on various batteries, the capabilities of battery systems for future planetary missions requiring flight durations of 3 to 5 yr or longer have not been studied. During most of the flight time, spacecraft power is delivered from solar panels or radioisotope thermoelectric generators; however, there are certain peak power requirements, such as planet encounter maneuvers, which require the use of storage batteries. The problem then is how to provide a battery system which maintains its coulombic efficiency after long periods of inactivity. There may also be various cycling requirements if the spacecraft is put into planetary orbit.

There are several approaches to the problem that can be investigated. One approach is to use automatically activated drycharged batteries similar to the systems used in missiles and rockets. Another approach is to store charged primary batteries such as the mercury-cadmium system. A third approach is to store a charged secondary battery system at some reduced temperature, then heat the battery to its operating temperature just prior to its required use. Other approaches are modifications of the third approach where the battery might be stored on a float charge or stored in the discharged state or even in the unformed state. This article discusses an investigation of the third approach. The AgO-Zn system is the prime system for investigation at present because of its high-energy density.

2. Test Program and Discussion

A storage optimization and design studies program is presently in progress at JPL and the battery test facilities at the Naval Ammunition Depot, Crane, Indiana, which will extend over a period of several years. Approximately 500 AgO-Zn cells with five different separator systems and four electrolyte concentrations are on test (Fig. 13). Another group of 100 cells with polyvinyl-alcohol and cellophane separator systems is being added to the program. When new separators are developed and become available, they will be evaluated. The cells were purchased from the various battery manufacturers, which adds an uncontrollable variable to the test matrix when attempting to compare the different separator systems; however, it should give a good measure to compare the state-of-the-art among the battery manufacturers. A group of 54 AgO-Cd cells is also on test (Fig. 13).

The basic test plan is to store the cells for various time periods at the temperatures noted in Fig. 13 and discharge the cells at 25°C to determine their capacity loss. An initial capacity check was made on all cells prior to storage. After storage in the charged state, the cells are given a second capacity check to measure the capacity loss during storage (referred to as residual capacity). The cells are given a third capacity check immediately after the second capacity check to determine if the original capacity can be restored (referred to as restorable capacity). The cells are then placed on a 24-h charge/discharge cycle (3 h discharge and 21 h charge) at 60% depth of discharge.

A summary of capacity data generated to date for this test program (approximately $\frac{1}{3}$ completed) is presented in Table 2. The average capacity of each of the nine different cell groups is given for different storage times and different storage temperatures as a percentage of average original cell capacity. By using average percentages of average initial capacities, it was possible to avoid the wide variation in the actual capacities of the cells of any given manufacturer type compared to its rated capacity as a factor and thus provide a more acceptable basis for comparing performance of these cells as reflected by their discharge capacities before and after their respective storage conditions. For example (Table 2), on the average the Delco, 25 A-h, AgO-Zn, fibrous sausaut casing separator, 50% KOH cells when discharged after storage at 25°C for 8 mo retained 83.7% of the average capacity of that cell group prior to storage. Furthermore, on the average 86.6% of the cell group's original capacity can be restored after storage by performing an additional cycle.

A summary of the cycle data generated to date is presented in Table 3. The average cycle life after storage of each of the nine different cell groups is given for different storage conditions. The nature of "average cycle life" should be kept in mind when evaluating the cycle performance of the different cell groups. Average cycle life does not directly reflect the performance of the poorest performing cell in a group, that cell which limits the performance of a battery. A thorough discussion of cycle life, which would require presenting in detail the rather extensive quantity of cycle life data generated on this test program, is beyond the scope of this article. However, a JPL technical report on this test program is forthcoming and will contain a detailed presentation of test results.

3. Conclusions

Since this test program is still in its initial phase, it is too early to make definite conclusions; however, the

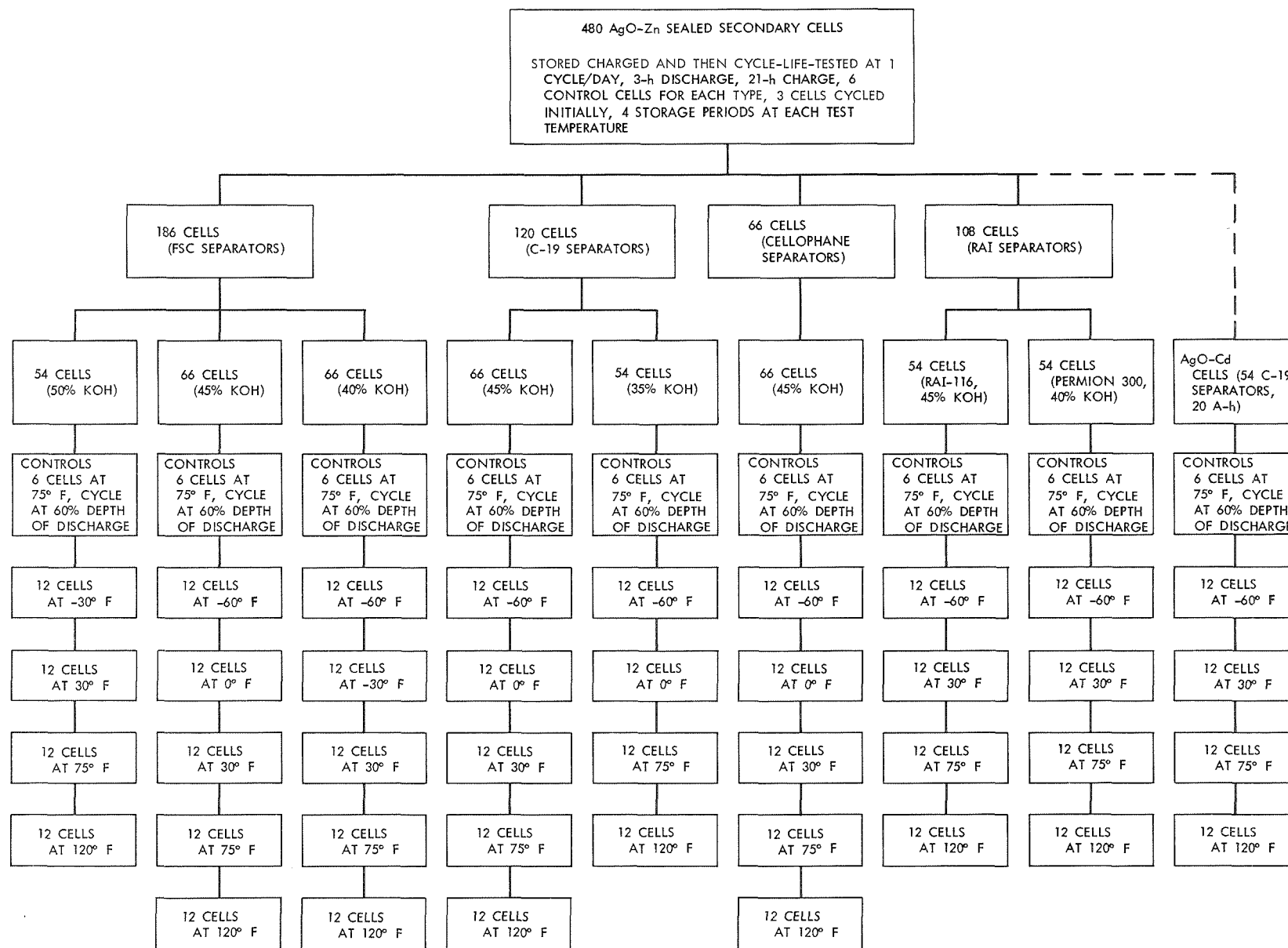


Fig. 13. Cell test plan

Table 2. Average cell capacity after storage for various times at selected temperatures

Cell ^a	Storage time, mo	Percentage of original capacity									
		-51°C		-35°C		0°C		25°C		49°C	
		Residual	Restorable	Residual	Restorable	Residual	Restorable	Residual	Restorable	Residual	Restorable
Delco, 25 A-h, AgO-Zn, FSC separator, 50% KOH	2										
	8			84.7	78.5	95.3	96.5	83.7	86.6	63.9	95.8 ^b
	16			95.0	91.1	93.6	100.3	74.0	80.6	26.9	29.4
Delco, 25 A-h, AgO-Zn, FSC separator, 45% KOH	2										
	8	77.6	82.7	78.3	85.7	93.6	93.7	78.5	82.5	57.0	0 ^b
	16	71.4	76.3	84.0	87.8	91.2	92.6	57.8	66.1	32.4	49.7
Delco, 25 A-h, AgO-Zn, FSC separator, 40% KOH	2									61.1	75.0
	8	92.7	109.1	91.7	104.2	92.7	99.0	80.1	82.8	39.0	41.6 ^b
	16										
Yardney, 25 A-h, AgO-Zn, C-19 separator, 45% KOH	2									47.1	80.1
	8	91.8	85.8			97.1	96.6	90.5	97.4		
	16	85.0	87.1			95.1	95.6				
Yardney, 25 A-h, AgO-Zn, C-19 separator, 35% KOH	2									34.0	43.8
	8	86.7	91.3					80.1	95.0		
	16	89.4	92.9					60.6	73.4		
ESB, 25 A-h, AgO-Zn, cellophane separator, 45% KOH	2									87.6	81.7
	8	118.4	115.0			115.1	102.3	112.9	80.6	71.2	65.3 ^b
	16	117.3	114.0			102.6	101.5	104.1	69.4		
ESB, 25 A-h, AgO-Zn, RAI-116 separator, 45% KOH	2									80.6	87.7
	8	138.0	92.6			88.1	69.2	91.1	84.1	53.2	0 ^b
	16	98.5	95.4			100.2	85.5	81.0	51.8		
Whittaker, 43 A-h, AgO-Zn, RAI-300 separator, 40% KOH	2									23.3	37.4
	8							29.2	63.5		
	16	36.0	98.1			0	91.3	0	19.3		
Yardney, 20 A-h, AgO-Cd, C-19 separator, 35% KOH	2									2.0	0
	8	81.3	98.9					54.4	56.1	0	0 ^b
	16	80.8	88.7					0.7	21.6		

^aFSC = fibrous sausaut casing, C-19 = proprietary modified cellophane, RAI-116 = modified polyethylene, RAI-300 = modified polyethylene.
^b4-mo storage.

Table 3. Cycle life after storage for various times at selected temperatures

Cell ^a	Cycle life average										
	Control cell	-51°C		-35°C		0°C		25°C		49°C	
		8 mo	16 mo	8 mo	16 mo	8 mo	16 mo	8 mo	16 mo	2 mo	4 mo
Delco, 25 A-h, AgO-Zn, FSC separator, 50% KOH	198	96		247	226	188	79	230	126	217	0
Delco, 25 A-h, AgO-Zn, FSC separator, 45% KOH	143	95	77	106	95	146	92 ^b	167	81	173	34
Delco, 25 A-h, AgO-Zn, FSC separator, 40% KOH	85	47		47		47		75		199	23
Yardney, 25 A-h, AgO-Zn, C-19 separator, 45% KOH	14	226	21			240	21		130 ^b	70	82 ^c
Yardney, 25 A-h, AgO-Zn, C-19 separator, 35% KOH	89	127	115			137	21	122	120	0	0
ESB, 25 A-h, AgO-Zn, cellophane separator, 45% KOH	24	24	30			100	84	90	47	105	79
ESB, 25 A-h, AgO-Zn, RAI-116 separator, 45% KOH	78	55	73			123	73	90	47	98	19
Whittaker, 43 A-h, AgO-Zn, RAI-300 separator, 40% KOH	49		10				13	40	0	5	0
Yardney, 20 A-h, AgO-Cd, C-19 separator, 35% KOH	232	40	153			215	163	68	0	0	0

^aFSC = fibrous sausage casing, C-19 = proprietary modified cellophane, RAI-116 = modified polyethylene, RAI-300 = modified polyethylene.
^b12-mo storage.
^c3-mo storage.

following trends are noted from data summarized in Tables 2 and 3:

- (1) AgO-Cd cells are not performing as well as the majority of the AgO-Zn cells.
- (2) Cells containing the RAI-300 separator material are performing poorly on stand and cycle life.
- (3) Most of the AgO-Zn cell groups have shown about equal capacity retention on stand thus far (the maximum stand period attained thus far is 16 mo).
- (4) There appears to be little difference between the performance of cells built with RAI-116 and cellophane separators. Most of the failures appear to be due to low capacity on cycling, which is thought to be caused by loss of negative plate activity.
- (5) Cells with high electrolyte concentration (50% KOH) yield the best cycle life.
- (6) Storage at the lower temperatures (0 to -51°C) is best for maintaining capacity on stand, as one would expect; however, the effect of storage temperature on capacity within the range 0 to -51°C appears to be slight.

F. Impedance of Silver-Zinc Battery Cells, W. L. Long

1. Introduction

The impedances of most types of battery cells have been studied for many years (Refs. 1-4). However, impedance data on silver-zinc battery cells has been extremely limited. This study measured the impedances of one model (No. 257, ESB, Inc.) of a 50 A-h sealed Ag-Zn cell under varying conditions of storage, charge, and discharge. This data may not apply to other types of Ag-Zn battery cells.

2. Instrumentation

All measurements were made with a commercial milliohmmeter (Keithley Instruments model 503) modified by placing 12,000-μF electrolytic capacitors in series with both the constant current leads and the voltage sensing leads. The instrument produced a 40-Hz square-wave constant current in one pair of leads and measured voltage with a synchronous vacuum-tube voltmeter with the other pair of leads. The indicating meter was calibrated in milliohms and ohms; the full scale ranged from 1 mΩ to 1 kΩ. Calibration checks with shunts between 0.005 to 1 Ω indicated that the instrument was accurate within ±2% of full scale.

3. Results

Impedance measurements of fully charged cells standing at room temperature and 100°F are shown in Figs. 14 and 15, respectively. Zero time in each case was at the completion of charge or recharge. All discharges were at

Table 4. Impedance change after 1-h discharge at 0.5 A

Impedance	Cell No. (ESB model 257) ^a		
	943	944	945
before discharge, Ω	0.041	1.85	1.40
24-h after discharge, Ω	0.055	2.45	2.04

^aAll cells completed 12-mo stand at 40°F before test; a 100% cycle was given cell No. 943 1 wk before test.

room temperature, 10-A constant current to 1.4 V. Recharges were in two steps: 2.0 A to 1.97 V, and 0.6 A to 1.97 V per cell. At both temperatures impedance increases with time, but the increase is greater at 100°F. Also, at 100°F, there is no difference in impedance increase between new cells and cells which have been cycled. The data points of the room temperature cells (Fig. 14) are grouped in two different areas, showing the higher impedance increase in the new cells.

Table 4 lists the impedances of three cells before and 24 h after a 1-h discharge at 0.5 A. The impedance of each cell increased.

Figures 16 and 17 show the impedance and voltage measurements recorded during battery cell charging and discharging, respectively. During charging, the impedance

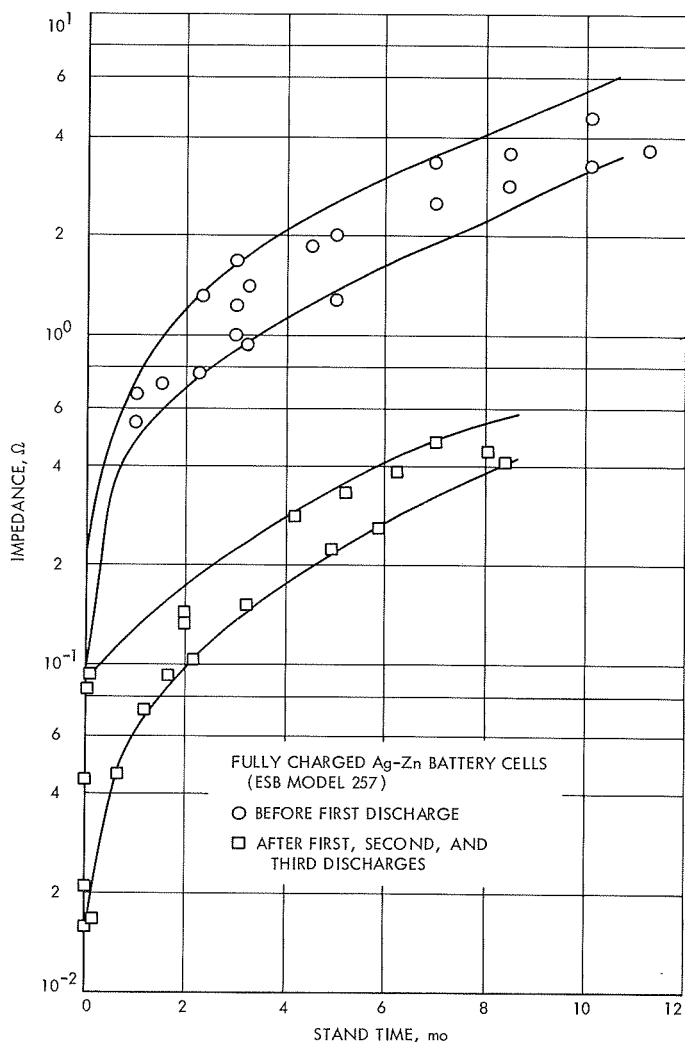


Fig. 14. Impedance vs stand time at room temperature

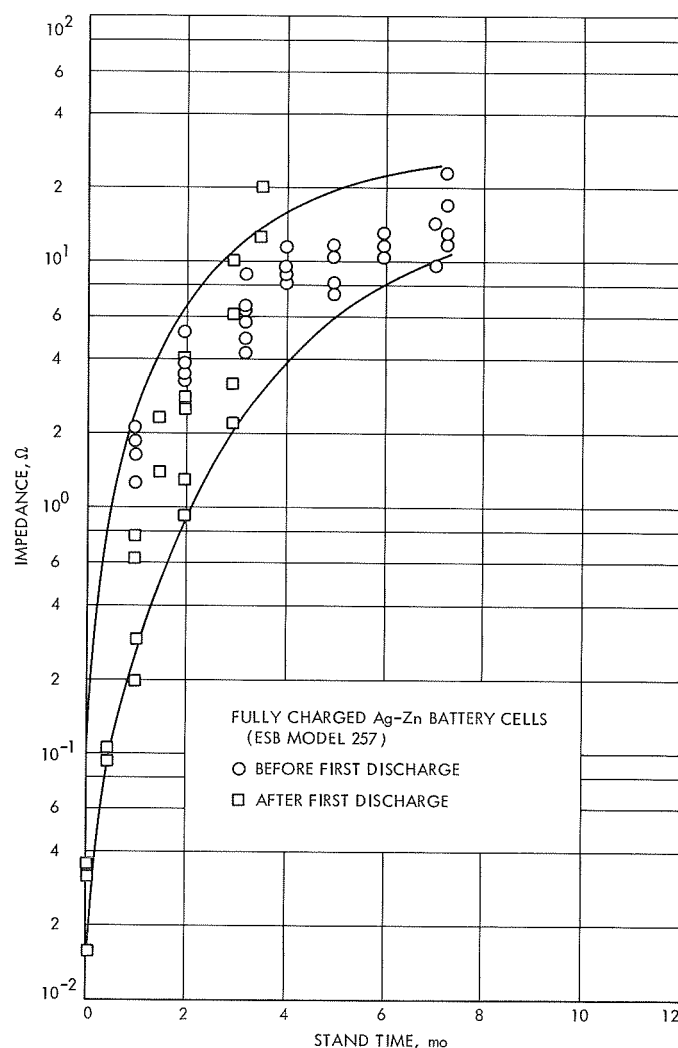


Fig. 15. Impedance vs stand time at 100°F

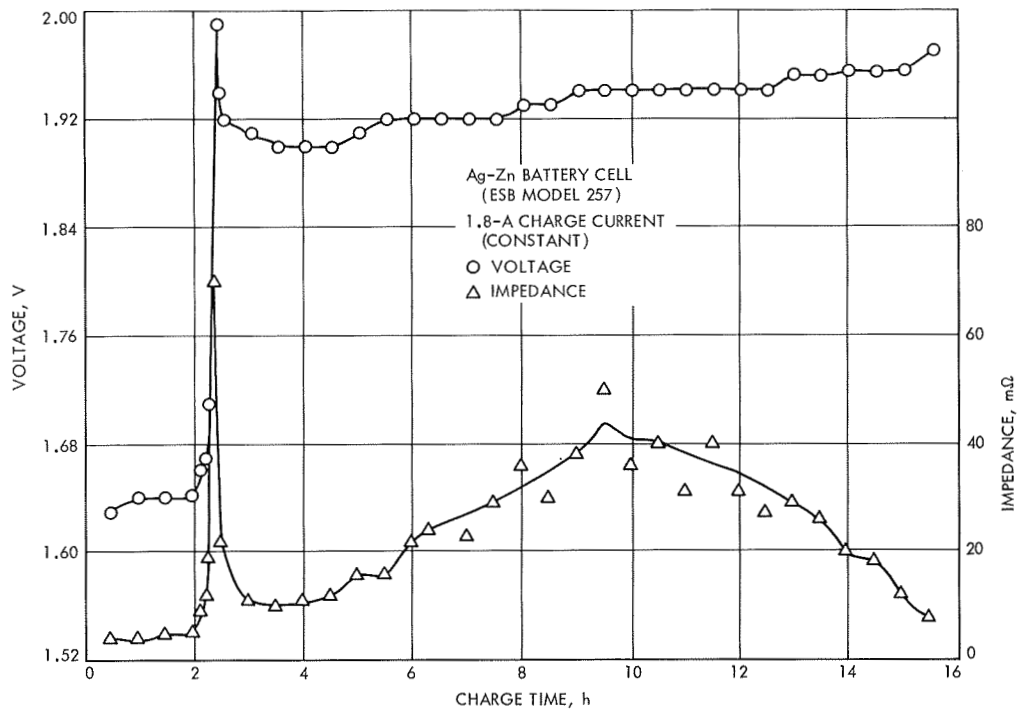


Fig. 16. Voltage and impedance during charge

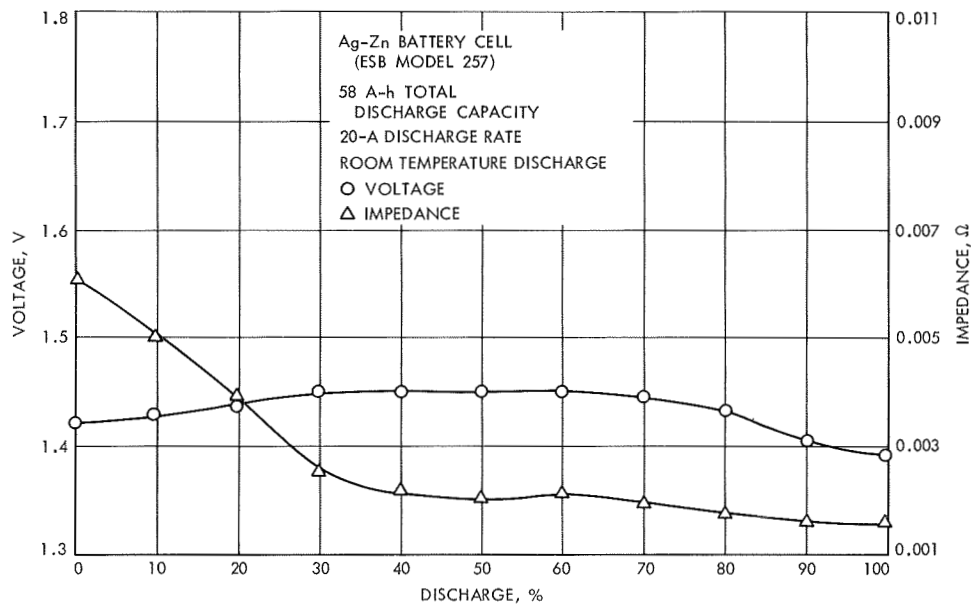


Fig. 17. Voltage and impedance during discharge

peaks at the same time as the voltage peak between the Ag_2O and AgO plateaus. During discharge at the 20-A rate, the impedance remains at a low value.

A summary of the impedance ranges is listed in Table 5. The overall range under all conditions tested is 0.002–20 Ω .

4. Discussion

The impedance changes are attributed to changing composition of the silver electrode. During charge, silver is oxidized in two steps: first to Ag_2O , then to AgO (Ref. 5). All three materials, Ag , Ag_2O , and AgO , may be present in the electrode at the same time. Table 6 lists the resistivities of these materials, showing an extremely wide range of values (Ref. 6). Possible explanations of the voltage peak in Fig. 16, which always appears during charging of the Ag-Zn cell, have been discussed by Butler in Ref. 7. Both the voltage and impedance peaks may be a result of a layer of high resistance Ag_2O completely covering the silver electrode. Then, as AgO starts to form, impedance and voltages become lower due to the smaller resistivity of AgO .

The low impedance during discharge (Fig. 17) may be due to silver metal being formed, resulting in a low resistance path for electrons. However, after only a short discharge (Table 4), the Ag reacts with AgO to form Ag_2O and increases the impedance. A fully discharged cell has a large percentage of silver, and retains its low impedance.

The increasing impedance of fully charged cells during storage (Figs. 14 and 15) is due to decomposition of

surface AgO to Ag_2O and oxygen. Higher temperatures increase the rate of decomposition.

5. Conclusions

The impedance of fully charged Ag-Zn cells increases with time on open-circuit stand. The increase is greater at 100°F than at room temperature. Cell impedance is below 0.01 Ω during high-rate discharge, and a fully discharged cell maintains this low impedance. Cell impedance may be raised or lowered by judicious choice of load current and time.

References

1. Letter Circular LC 965, National Bureau of Standards, Washington, D.C., Nov. 15, 1949.
2. Brodd, R. J., and De Wane, H. J., "Impedance of Sealed Nickel-Cadmium Dry Cells," *Electrochem. Technol.*, Vol. 3, p. 12, 1965.
3. Clark, R. P., and Zaffery, E. D., "Instrument for Continuous Measurement of Internal Resistance of Voltaic Cells During Discharge," *Rev. Sci. Instr.*, Vol. 38, p. 492, 1967.
4. Vinal, G. W., *Storage Batteries*, 4th Edition, p. 328. John Wiley & Sons, Inc., New York, N.Y., 1955.
5. Bauer, P., *Batteries for Space Power Systems*, NASA SP-172, p. 147. National Aeronautics and Space Administration, Washington, D.C., 1968.
6. Tvarusko, A., "The Electric Resistivity of AgO ," *J. Electrochem. Soc.*, Vol. 115, p. 1106, 1968.
7. Butler, E. A., *Studies of Reaction Geometry in Oxidation and Reduction of the Alkaline Silver Electrode*, Final Report Contract 951157, JPL Reorder No. 66-259. Brigham Young University, Provo, Utah, Apr. 1966.

G. Heat Sterilizable Battery Development, R. Lutwack

JPL subcontract 951296 with ESB, Inc., a program to develop sterilizable Ag-Zn cells, is divided into two main parts, one for sterilizable cells and the other for sterilizable, high-impact resistant cells. The first part contains developments of a 70 A-h primary cell capable of four 100% cycles, a 10 A-h primary cell capable of four 100% cycles, a 25 A-h secondary cell capable of ninety 60% cycles, and a 48 A-h secondary cell capable of four hundred 60% cycles. The other part comprises developments of a 5 A-h primary cell capable of four 100% cycles and a 25 A-h secondary cell capable of ninety 60% cycles after sustaining a high-impact shock of 4000 g. The status of these developments follows.

Data have been obtained for sterilized and unsterilized 70 A-h primary cells of a modified design. The input charge for all cells was very good, varying from 0.38 to 0.44 A-h/g Ag ; the unsterilized cells accepted an average

Table 5. Impedance ranges (Mariner Mars 1969 Ag-Zn cells)

Condition	Impedance range, Ω
Room-temperature storage-charged before first discharge	0.03–5.0
after first discharge	0.02–0.6
100°F storage-charged	0.02–20.0
During 1.8-A charge	0.002–0.07
During 20-A discharge	0.002–0.006

Table 6. Resistivity of silver and its oxides

Material	Resistivity, $\Omega\text{-cm}$
Ag	1.6×10^{-6}
Ag_2O	10^8
AgO	10

5.5% more charge than did the sterilized cells. The formation discharges were also very good, varying from 0.34 to 0.42 A-h/g Ag; the sterilized cells delivered an average of 8% less than did the unsterilized cells. The unsterilized cells have completed three additional cycles with a minimum discharge of 0.38 A-h/g Ag. The sterilized cells have been fabricated into a six-cell monoblock, which is to be vibration-tested, after which additional electrical testing will be done. The conclusion from these data is that there is now available a high-energy density primary 70 A-h cell capable of providing satisfactory electrical performance of four 100% cycles after heat sterilization; it has yet to be demonstrated that this cell will withstand vibration testing.

Additional data have been obtained for the 25 A-h secondary cell (90 cycles at 60% depth of discharge), which is the prototype of the cells being evaluated for the *Viking* lander. Several groups of cells of different designs have been electrically cycled after heat sterilization. In the group with the best performance, no failures occurred for 160 cycles, at which point 16.6 A-h was the delivered capacity. Conclusions derived from the performance of these cells are:

- (1) The use of a positive plate absorber is beneficial for cycle life.
- (2) The use of a positive plate separator wrap leads to longer cycle life than does a negative plate separator wrap.
- (3) Electrical cycling before heat sterilization reduces cycle life.

This reduction is as much as 54% in the group with the best performance. (However, there are data from other cells that show that this loss is a result of the condition that this group of cells was subjected to 8 cycles before sterilization.) Analyses of failed cells have led to these conclusions:

- (1) The use of absorbers results in a pattern of random erosion of the negative plate active material rather than the more catastrophic bell-shaped profile which occurs in the absence of an absorber. The random pattern appears to be caused by the capability of the absorber to maintain electrolyte solution over the entire surface of the negative plate.
- (2) The procedure of cycling a cell before sterilization leads to a larger amount of silver trapped in the separator as compared with the case when there is no pre-sterilization cycling.

The design of impact-resistant cells incorporates a Zr sheet in the positive plate and an etched Ag sheet in the

negative plate. Satisfactory discharge efficiencies have been obtained for cells with the Zr-positive plates; for the first two cycles, the efficiencies ranged from 0.34 to 0.36 A-h/g Ag. Analyses of cells with Zn plates fabricated with etched Ag sheets determined that the Ag sheets were in a greatly weakened and brittle condition. In a study of this problem, which incorporated x-ray diffraction and chemical analysis as well as cell measurements, it has been shown that the amalgamation process alone does not cause this effect; the reduction of strength, however, occurs if the plate is subjected to the heat treatment of sterilization or of the sintering process. For example, sintering at 325°C for 1 h results in a reduction in yield strength of 72%. The x-ray diffraction data were interpreted as showing that there is no correlation between the brittleness of the grid and the changes in the metallic phases. The investigation of this phenomenon is continuing, and other materials are being examined as substitutes for the etched Ag plate.

A study of the variables in the design of an extended cycle life cell (400 cycles at 60% depth of discharge) is being done using a factorial experiment. No significant parameter has been identified yet.

H. Performance Testing of the SNAP-27 Thermoelectric Generator, G. Stapfer

1. Introduction

The initial performance testing of a SNAP-27 thermoelectric generator at JPL is described. The generator, designated Mod-15, is the flight hardware qualification generator for the ALSEP program.² Generator evaluation tests were performed at three nominal power input levels in addition to a low power input point, which served as a reference for start-up procedures.

The three levels of power input were: 1500, 1450, and 1400 W. This article describes the performance characteristics of the generator at those three input power levels. It briefly describes the generator, the test equipment, instrumentation, and the vacuum system utilized for the tests. It also contains curves and data, comparing the power output and efficiency of the generator at the various power input levels.

2. Generator Description

The SNAP-27 Mod-15 thermoelectric generator was developed by the General Electric Company. Major sub-contractors were the Solar Division of the International

²ALSEP = Apollo lunar surface experiments package.

Harvester Company and the 3M company. The generator is on extended loan to JPL from the Atomic Energy Commission for the purpose of performance and life testing.

The generator is 16 in. in diameter, 18 in. high, and weighs about 27 lb. The thermoelectric couples are manufactured from 3P-3N lead-telluride material. Thermal insulation, MIN-K-1301, is packed into the voids between the internal components of the generator. Argon gas, at a pressure of 25 psia provides the inert atmosphere required by the lead-telluride thermocouples. The generator's waste heat is dissipated to the surroundings by eight fins attached to the cold frame.

Figure 18 shows the generator installed in the test vacuum chamber. Both cold- and hot-frame temperatures are monitored by Chromel-Alumel thermocouples. The thermocouples which measure the hot-frame temperature are spotwelded directly to the inner casing. The generator's cold frame, or fin root temperatures, are monitored with thermocouples which are located near the base of the heat radiation fins.

A tungsten heater is used to simulate the isotope fuel capsule. This heater is inclosed in a sheath of the same dimensions and thermal characteristics as the actual fuel capsule.

3. Test Equipment

The generator is designed to be operated either in a vacuum environment or in air. However, special precautionary measures must be taken for air operation; JPL elected to perform the generator tests under vacuum conditions. The vacuum chamber used is 3 ft in diameter by 3 ft long. A VacIon pump is mounted in each of the two end domes. In this arrangement, the pumps constitute an integral part of the vacuum chamber, thereby increasing the actual pumping efficiency. Additional pumping capacity is added to the basic 600-l/s VacIon pump by installing titanium sublimation pumps in each of the two domes. This allows the generator to be operated at pressures of less than 1×10^{-7} torr. The entire double-walled stainless-steel vacuum system is water-cooled to maintain the inside walls of the chamber at ambient temperature.

The input and output power connections to the generator, as well as all the instrumentation, are brought out of the chamber through metal-to-ceramic feedthroughs. The output cable of the generator was cut approximately 6 in. from the generator to allow generator voltage measurements to be recorded without the voltage drops

normally associated with the original flight hardware interconnection cable.

The circuitry which controls the input and output power of the generator is basically an identical system to that being used to operate the SNAP-19, SNAP-11, and thermoelectric modules presently being tested by the JPL Nuclear Power Sources Group. The system will maintain the input power to the generator constant at $\pm 1\%$. An electronic load maintains the output voltage of the generator constant (± 10 mV) and is adjustable from open-circuit voltage to near short-circuit current. The generator is thermally protected by an over-temperature guard which reduces the input power to 50% of its original value should the hot-frame temperature exceed a predetermined limit. All power to the generator will be shut off should the vacuum pumps fail.

Voltage, current, and temperature measurements are made with a digital voltmeter (DVM) in addition to which panel meters are used for visual observations. An electronic reference junction provides a 0°C reference for all thermocouple measurements.

To insure an accurate open-circuit voltage reading, a time delay circuit will automatically lock the DVM a few milliseconds after the load circuit is interrupted. Both input and output power measurements are recorded by an automatic data acquisition system. Continuous records of these parameters, as well as the temperatures, are made at hourly intervals.

4. Generator Tests

The Nuclear Systems Operations Group of the General Electric Company performed the initial performance tests of the generator (Ref. 1). During these tests, approximately 500 h of generator operation time was recorded. The Bendix Corporation subsequently performed overall system tests on the complete ALSEP package, including the Mod-15 generator. During these tests, about 600 h of generator operation were recorded.

During the initial start-up of the SNAP-27 generator at JPL, very heavy outgassing of the generator was observed. It is presumed that this was due to the argon atmosphere in which the generator was stored during shipment.

For the purpose of an operational checkpoint, the generator was stabilized at a power input of 1000 W. Performance data of the generator, as well as of the test equipment, were established at this input power level. Both

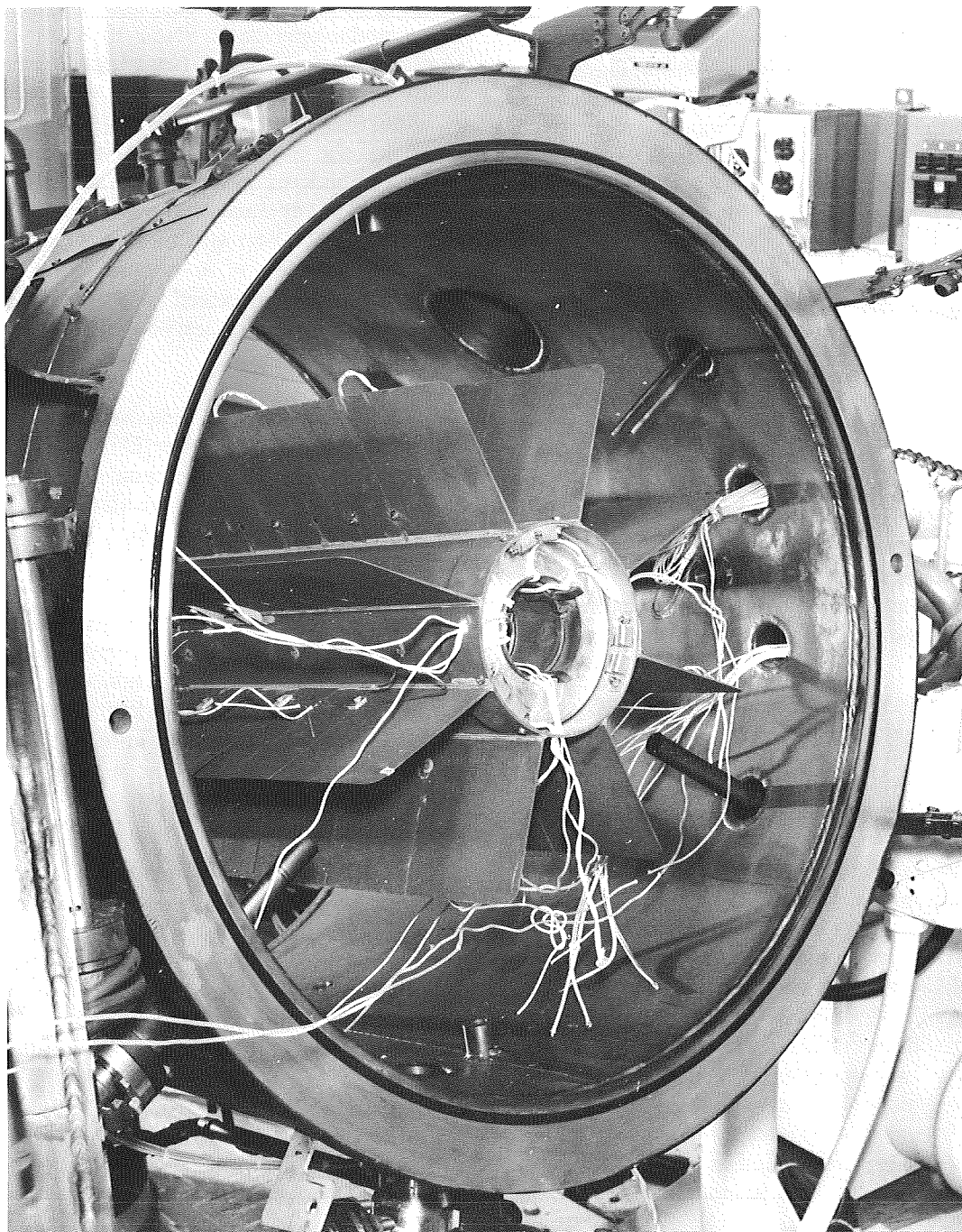


Fig. 18. SNAP-27 generator installed in test chamber

data are useful for future start-up operations since they constitute valid data points at a very low power input level. After the generator and the test equipment were confirmed to be operating satisfactorily, the input power level to the generator was increased to the performance evaluation test level.

A complete characteristic performance map of the generator was recorded at three different input power levels: 1500, 1450, and 1400 W. For each data point, the power input to the generator was maintained at $\pm 1\%$, while the output voltage was held constant (± 10 mV) at a fixed voltage. After complete thermal equilibrium of the generator was achieved (nominally 4–8 h) at the desired point, all pertinent performance parameters were recorded. A dynamic volt-ampere curve, characteristic of each steady-state operating point, was then plotted by an *x-y* recorder.

The performance characteristic of the generator at a given input power level was thus established by at least seven steady-state operating points. Figure 19 shows the performance of the generator at an input power of 1500 W. A maximum power output of 74.5 W was achieved at an

output voltage of 17 V. Figure 19 shows that the cold-frame temperature of the generator remained essentially constant at 200°C, while the hot-frame temperature varied from 525°C near short-circuit current to 605°C at an output voltage of 28 V. All temperatures plotted in Fig. 19 are the average temperatures of the measured cold- or hot-frame thermocouple readings. At 1450-W input power to the generator, a maximum power output of 71.2 W at hot- and cold-frame temperatures of 550 and 197°C, respectively, was observed. At an input power of 1400 W, the maximum power output of 65 W occurred at an output voltage of 16 V and at hot- and cold-frame temperatures of 535 and 193°C, respectively.

The generator output current and output power at the three input power levels are compared in Fig. 20 as a function of output voltage. As is expected, the output voltage, corresponding to the maximum output power point, shifts slightly as a function of input power. It is of interest to note that although the input power is changed at identical increments of 50 W, the output power increase between 1450 and 1500 W of input power is smaller than the output power change between 1400 and 1450 W of input power.

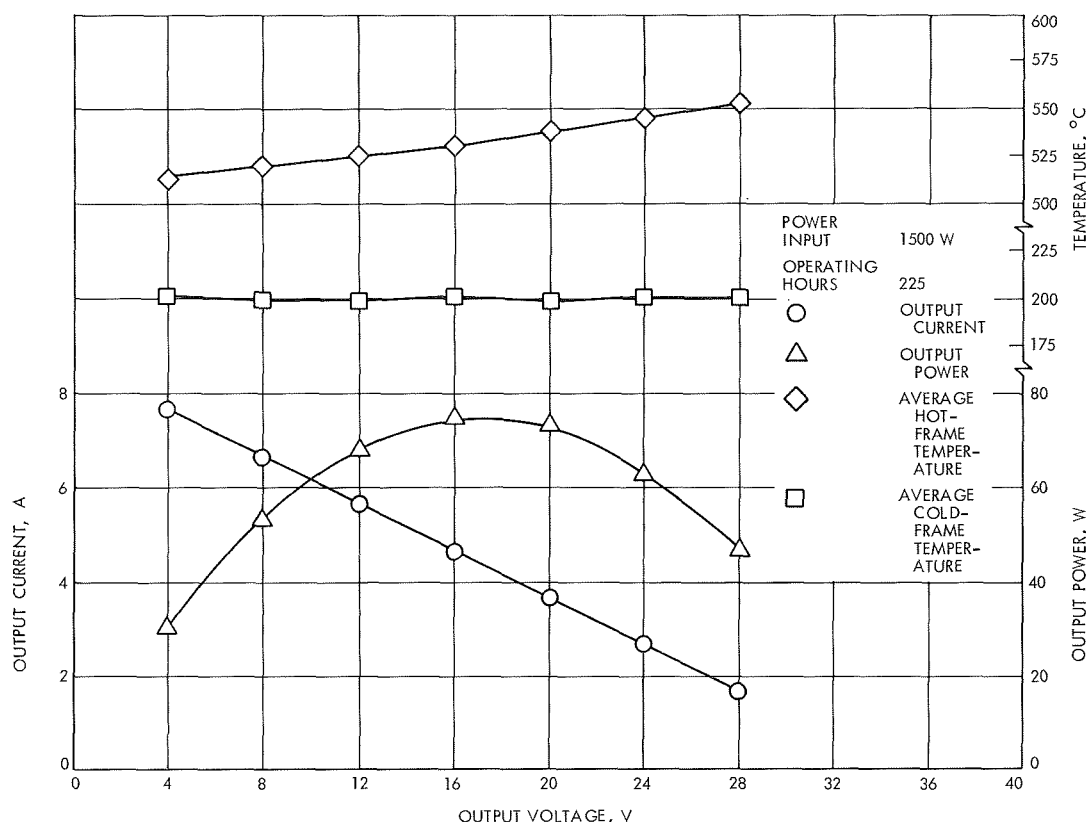


Fig. 19. SNAP-27 generator performance at 1500 W

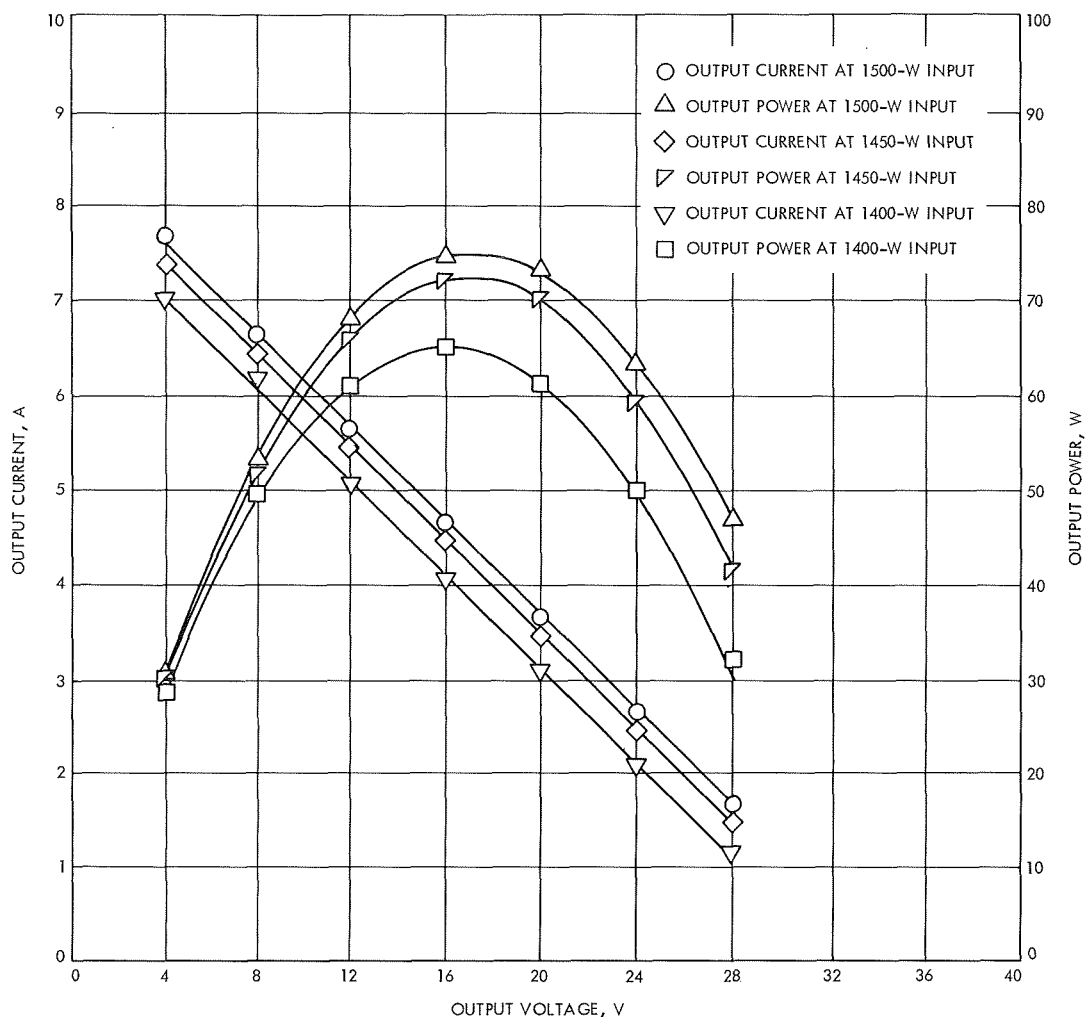


Fig. 20. SNAP-27 generator overall performance

5. Conclusion

The electrical performance of the generator at JPL was found to be in fair agreement with the initial performance data obtained by the General Electric Corporation (maximum power output at GE was 73.0 W; maximum power output at JPL was 74.5 W). The higher power output observed at JPL was primarily due to the shorter output cable and slightly different operating temperatures.

The generator is now continuing to operate at an input power of 1500 W and at an output voltage of 16 V for long-term endurance and life tests. The performance of the generator is continuously monitored, and the power output is plotted as a function of time on a biweekly basis. At intervals of approximately 2000 h of operation, a volt-ampere characteristic of the generator is recorded at the 1500-W input power level. The generator is also available

for the performance of special tests which may be required for future analysis.

Reference

1. SNAP-27 Generator Assembly Model-15 Acceptance Tests, Revision A, Test Report ANSO, Document 6300-258. General Electric Co., Valley Forge, Pa., Sep. 1967.

I. The Application of Bench Tests in the Development of Heat Sterilizable Battery Separators, W. von Hartmann

1. Introduction

Bench tests are used at the Jet Propulsion Laboratory in the development of heat sterilizable battery separators. The objectives of these tests are to provide criteria for the

quality control of the currently used materials, to test materials, and to support in-house electrochemical cell development. All of the efforts to date have been aimed at the silver oxide-zinc cell.

A material under investigation may be subjected to one or more of the following:

- (1) Sterilizability tests.
- (2) Area-resistance tests.
- (3) Tensile strength tests.
- (4) Dimensional change tests.
- (5) Silver oxide diffusion tests.
- (6) Crosslink density tests.
- (7) Spectrophotometric examination.
- (8) Various chemical tests.
- (9) In-cell tests.

Most materials are not subjected to all of the aforementioned tests. Frequently, it is found that one key test rules out the material in question.

In Table 7 are listed some of the materials tested to date. Since the material made from polyethylene, which has been modified by grafting and crosslinking, has been most successful, data obtained during tests on this material will be used to illustrate this work.

2. Sterilizability

The sterilizability of a material is usually the first quality to be determined. In this test, samples from a material

Table 7. Materials tested

Material	Source
Polybenzimidazoles	Whittaker Corp.
Ethylene-acrylic acid copolymer	Monsanto Research Corp.
Zirconium oxide and binder on polypropylene mesh	Westinghouse Electric Corp.
2-vinylpyridine-methylmethacrylate copolymer	Monsanto Research Corp.
Modified polyethylene	RAI Research Corp.
Cellophane	E. I. du Pont de Nemours & Co., Inc.
Microporous Teflon	E. I. du Pont de Nemours & Co., Inc.
Modified polyethylenes	Southwest Research Institute
Fibrous sausage casing	Union Carbide Corp.

are covered with a 40% aqueous solution of KOH to simulate the battery electrolyte, are sterilized at 135°C for 120 h, and are then examined for evidences of chemical attack and physical changes. In the absence of obvious damage, an area resistance determination is usually the next test to be made. In addition, other tests, which are described below, are used frequently to determine the effect of sterilization on the material.

3. Area Resistance

The resistance measurement of separators is used in the separator and cell development efforts as well as for quality control purposes. The fixture employed in this measurement is shown in Fig. 21, the sample under test being placed between the cell halves which are filled with electrolyte solution. The measurement is made with an ac bridge. The resistance is multiplied by the area of the separator exposed in the cell to yield a value of area resistance in $\Omega\text{-in.}^2$ To free the result from the uncertainty inherent in a thickness measurement on a wet and compressible material, the thickness of the material is not incorporated into the result but is reported separately. As an illustration of the use of the resistance measurement for quality control, the results obtained from a polyethylene-type separator by sampling 80 continuous ft of material 50 times are shown in Fig. 22. The histogram shows good uniformity and a relatively narrow spread in values.

4. Tensile Strength

Properties such as tensile strength and elongation are measured even though it is difficult to establish meaningful criteria for separators in terms of these qualities. In the case of the materials which are very elastic, the tensile strength measurement has little meaning. However, these measurements, which are made in air or in 40% KOH solution and at room temperature as well as at 135°C, can

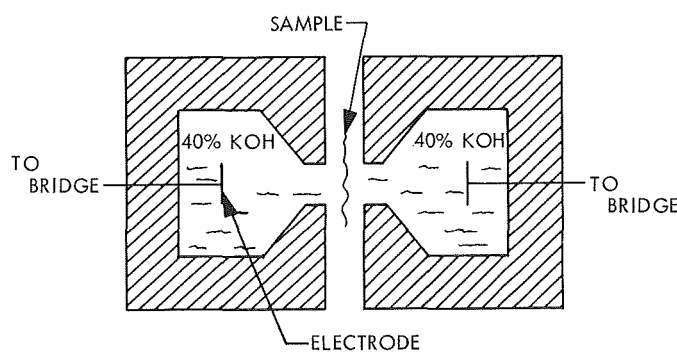


Fig. 21. Resistivity cell

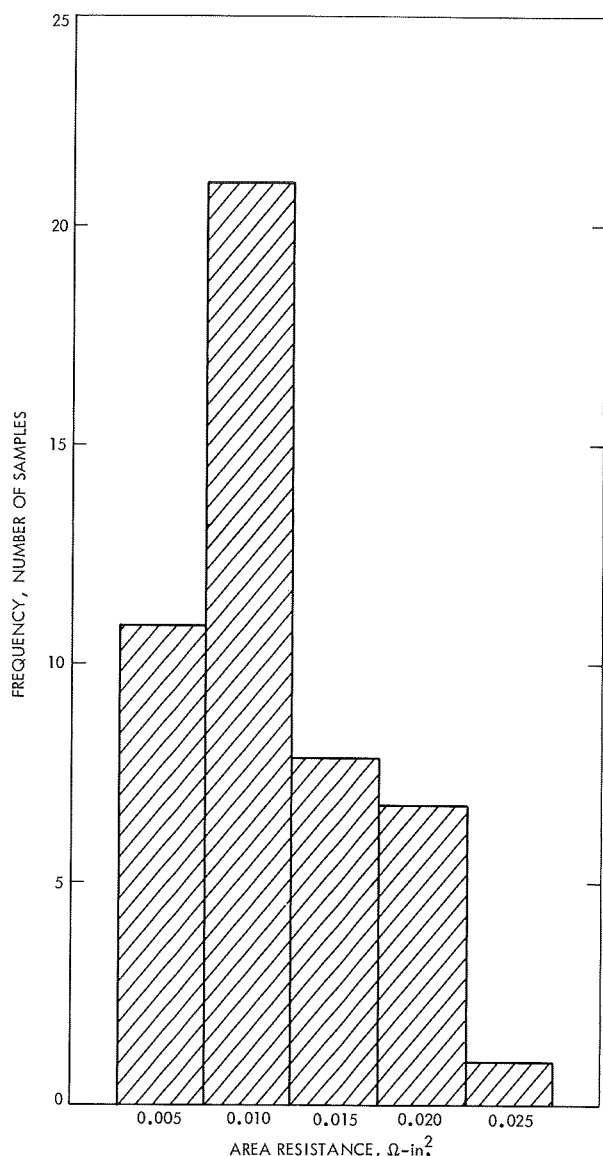


Fig. 22. Distribution of area resistance values for polyethylene-type separator

be used to determine whether structural degradation occurs during sterilization. The results have been used to establish manufacturing specifications. As an example, the polyethylene separator material typically has a tensile strength of about 800 lb/in.² and an elongation which exceeds 100%.

5. Dimensional Change

The swelling of separator materials as the result of electrolyte absorption is well known. In addition, sterilization causes irreversible dimensional changes. In the case of the polyethylene-type separator and perhaps other mate-

rials, this is caused by the relaxation at elevated temperatures of the stresses incorporated into these materials at the time of manufacture. Dimensional changes are determined by measuring samples between heavy glass plates, to assure flatness, first in the dry state, then following soaking in electrolyte solution, and finally after sterilization. The data are tabulated as percent dimensional change based on dry dimensions. In Table 8 are shown some representative data of dimensional change measurements made on a polyethylene-type separator. In comparing values of the soaked sample with those for the sterilized sample, it can be seen that considerable shrinkage occurs in the direction along the roll during sterilization. This is the result of the stress relaxation mentioned earlier. A comparison between the values for the crosslinked and uncrosslinked samples in this table shows that crosslinking reduces dimensional change in the material.

Table 8. Dimensional stability of polyethylene-type separator

Test	Soaked in KOH, % of dry	Sterilized, % of dry
Not crosslinked:		
Along roll	+ 12	+ 4
Across Roll	+ 11	+ 11
Crosslinked:		
Along roll	+ 8.5	- 2.3
Across roll	+ 8.8	+ 6.5

6. Silver Oxide Diffusion

Silver oxide is slightly soluble in the electrolyte solution, diffuses through the separator, and reaches the zinc electrode where the resulting local cell action would discharge that electrode. In addition, the silver oxide can oxidize separator materials. The silver oxide is reduced in the process and remains in the membrane as metallic silver. When enough silver has been deposited by this process, the electrodes short through the separator and the cell is discharged internally. This process occurs frequently enough to make silver oxide migration a major failure mode in the silver oxide-zinc cell. The measurement of the silver oxide diffusion rates is being used in a study of the effects of membrane modifications on the diffusion. A two-compartment diffusion cell is used (Fig. 23). One compartment is filled with electrolyte solution while the other contains an electrolyte solution saturated with silver oxide to which has been added a small quantity of radioactive silver 110 isotope. The rate of silver oxide diffusion and the amount of silver/silver oxide deposited in the membrane are determined by standard radioisotope methods.

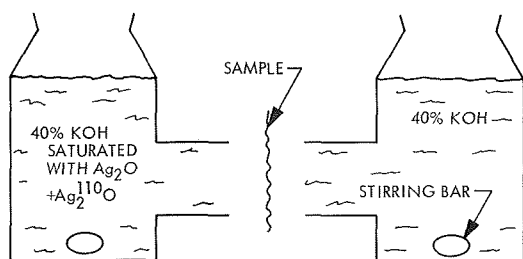


Fig. 23. Diffusion cell

Some typical diffusion data are given in Table 9. These data are being used to find materials with minimum silver oxide permeability and, in the case of polyethylene-type separators, to study whether the crosslink density has any effect on permeability. An additional study is being made of the relation of separator resistance to silver oxide diffusion, as a function of silver loading and the use of additives.

Table 9. Silver oxide diffusion through battery separators

Sample	Diffusion, g/h-in. ²
152 GX	1.95×10^{-2}
	2.19×10^{-2}
	1.80×10^{-2}
Cellophane	0.48×10^{-2}
	0.65×10^{-2}
P 1770	1.31×10^{-2}
146 GX	2.88×10^{-2}

7. Crosslink Density

The relationship of the extent of crosslinking of polyethylene separator materials to the properties of resistivity, tensile strength, and diffusion rate is being studied. For these materials, the degree of crosslinking is determined via a tensile measurement (SPS 37-54, Vol. III, pp. 100-104). In this procedure, a sample is soaked in a good polyethylene solvent such as alphachloronaphthalene at 125°C. The sample is pulled while submerged in the solvent at the elevated temperature. Under these circumstances the sample swells severely, and only the crosslinked chains contribute measurably to the tensile strength. From the tensile strength and elongation measurements obtained, and a knowledge of the weight of the sample, a value for the crosslink density can be calculated. For the polyethylene-type separator, a typical value is 10^{-3} moles crosslink chains/cm³. Attempts to correlate these values with other characteristics of the material are continuing.

8. Spectrophotometric Examination

IR spectra are being used to distinguish between the acid and the salt form of the acrylic and methacrylic acid groups which were grafted in the polyethylene-based membranes. The spectra also serve to identify methyl and methylene groups in the base material and can be used to measure accurately the thickness of the base material. UV spectra are being used to determine the concentration of the benzene moiety in membranes in which divinylbenzene is the crosslinking agent.

9. Various Chemical Tests

Many of the properties of the polyethylene-type separator depend upon the kind of polyethylene used as the starting material. Consequently, each shipment of polyethylene is carefully examined to assure uniformity. The following parameters are being determined: molecular weight, molecular weight distribution, density, melting point, crystallinity, degree of branching, and the presence of slip additives, fillers, or other impurities (the procedures and some results have been described in SPS 37-50, Vol. III, pp. 166-169). A set of typical values for the polyethylene used is shown in Table 10.

Table 10. Properties of the polyethylene used in the manufacture of battery separators

Parameter	Value
Density at 22°C	0.925 g/cm ³
Melting point	107°C
Crystallinity	50%
Degree of branching	2 methyl groups per 100 carbon atoms
Solubility	100% in alphachloronaphthalene above 107°C
Unsaturation	0.12 mM/g of pendent vinylidene groups
Molecular weight	48,200
Molecular number	15,000
Molecular weight/molecular number	3.2

10. In-Cell Tests

A given material could pass all the tests described so far and still not be usable as a battery separator. Therefore, in-cell testing is performed to establish the characteristics of the material as deduced from the performance of AgO-Zn cells. These tests are undoubtedly the most important ones since they establish the required physical

and chemical compatibility of the material with the cell environment. Nevertheless, since it is very difficult to interpret cell results because of the large number of variables operating in a cell, considerable care must be taken in devising the tests and interpreting the results to be sure that the tests describe the dependence of the cell performance on the separator.

A typical cell for these tests is built by wrapping four zinc oxide electrodes with three to five layers of separator material, inserting three silver electrodes so that each silver electrode is faced by zinc oxide electrodes on both sides, and placing the resulting cell pack into a cell case. Electrolyte solution, 40% aqueous KOH, is added and the case is sealed. The finished cell is sterilized for 120 h at 135°C. All cells are cycled four times at constant current with voltage cutoffs of 1.2 and 2.0 V. Some of these cells are charged once again and placed on an open-circuit charged stand. The open-circuit voltages are monitored daily, a cell falling below 1.84 V being considered to have shorted.

This stand test has been used for heat-sterilized cells for only a short time so that no failures have occurred to date. In other tests, the cells are simply cycled under the same regime until they no longer accept or hold a charge. Some cells have completed as many as 120 full cycles, while others failed after 40 cycles. Cells which have failed are disassembled to determine, whenever possible, the reason for failure. The majority of failures found in cells which were cycled deeply seems to be caused by Ag shorting caused by the reduction of silver oxide in the separator, thus forming a metallic path between the electrodes.

11. Conclusion

The bench tests mentioned here have proven to be useful in qualifying a separator material for a specific application. The list of tests is by no means exhaustive and other tests will be used as the need arises. Although the polyethylene-based battery separator material has been found to be acceptable for heat sterilizable batteries, efforts are being made to improve the material. In particular, a task devoted to increasing its resistance to silver oxide diffusion is being emphasized.

J. Long-Term Life Test of the Spare

Mariner Venus 67 Power Subsystem Hardware, A. Krug

The long-term life test of the spare *Mariner Venus 67* power subsystem hardware will provide a measure of

the power system designs and screened components capability to survive missions to the outer planets. The test purpose and method is described in SPS 37-57, Vol. III, p. 76.

The contract negotiations were satisfactorily completed with the selected contractor, and at that time the initial duration of the life test was extended from 12 to 17 mo. The purpose of this extension was to obtain a time period equal to a nominal trajectory between earth and Jupiter.

Since the first of October, the contractor has been preparing the required test procedure and setting up the environmental equipment preparatory to the start of the long-term life test.

The operational support equipment (OSE) and flight hardware, which were checked out at JPL prior to shipment, were operating normally. A checkout of the equipment at the contractors after completion of the setup also indicated normal operation, with only a couple of minor anomalies in the OSE, which were quickly corrected.

It is presently planned to start the life test in the first week of December 1969, which would place the completion of the 17-mo period in May 1971. Data will be taken on the overall operation at 1-wk intervals and submitted to JPL in a monthly report. This will allow continuous monitoring of the test results and the capability to arrive at some conclusions about the results prior to the completion of the 17-mo period.

K. Interactions Between Radiation Fields From RTGs and Scientific Experiments on Spacecraft, C. G. Miller and V. C. Truscello

1. Introduction

For missions to the outer planets, the most promising source of onboard spacecraft power is the radioisotope thermoelectric generator (RTG). At the present state of the art, the RTG can produce approximately 1 W/lb. At this rating, it matches present solar panels in output at distances of 3 AU from the sun, and retains this rating at Jupiter and Saturn distances where solar panels fall to one-half and one-tenth the RTG output, respectively.

The distinctive characteristics of an RTG in a spacecraft application that are considered in this article are the result of the incidental but ubiquitous neutron and gamma radiation fields accompanying the RTG. The major

objection to the use of an RTG is the potential interference with the scientific equipment on board the spacecraft. This is due to the fact that the stray radiations from the RTG may affect a particle experiment in a manner that cannot easily be distinguished from the desired signal. This article examines the various particle experiments and components of experiments that have been proposed, to see which are amenable to reduction of such interference. Among the measures that may be taken are the use of coincidence techniques, correction for background values, and appropriate additional shielding. The use of any such measures exacts a penalty in spacecraft weight, complexity, or reliability, so every experiment and the measures needed to reduce RTG interference must be optimized individually.

The evaluation is complicated somewhat by the fact that the RTG power output and radiation spectra change considerably with time as decay products build up. The design must be such that worst-case situations are taken into account.

This article considers the shielding required, the spatial deployment, and experiment modifications needed to assure compatibility of RTG power sources and scientific experiments on the same spacecraft.

2. Approach to the Problem

The overall approach taken to evaluate the instrument sensitivity and develop the shielding requirements follows:

- (1) Instruments which were suspected of being affected by the RTG radiation were selected for evaluation.
- (2) A thorough study of the operation and description of these instruments was made to determine sensitive components and dimensions of the sensitive volumes. The minimum charged particle counting rate of each detector was also established.
- (3) Since the sensitivity of the detector to RTG radiation (gamma and neutron) varies with the energy of the radiation, an accurate spectral description of the emitted radiation is necessary. The following spectral information was determined in the form of a 20-group structure:
 - (a) The magnitude and description of both the neutrons and gammas emitted by the radiation source at various points outside of the RTG envelope.
 - (b) The reduction in magnitude and change in spectral distribution of the gamma radiation as

a result of placing various thicknesses of tungsten shield between the RTG and detector.

- (c) The magnitude and spectrum of the gamma radiation produced by the interaction of the neutrons with the shield material.
- (4) A suitable analytical model to predict the effects of the radiation spectra of item (3) upon the following sensitive components was developed:
 - (a) Solid-state detectors.
 - (b) Geiger-Muller tubes.
 - (c) Continuous channel multipliers.
- (5) An acceptable value of spurious counting was established.
- (6) By using the data from the analytical model and the established values of allowable interferences, the amount of shielding required for each of the experiments consisting of one or more of the above radiation sensitive components was determined.

3. Instrument Susceptibility

Spacecraft instrumentation may be broadly classified according to the species of particle being examined:

- (1) Positive particle detectors.
- (2) Negative particle detectors.
- (3) Neutron detectors.
- (4) Radio frequency measurements.
- (5) Optical and near-optical frequency measurements.
- (6) X-ray and gamma-ray measurements.
- (7) Magnetic field measurements.

These seven general types of instruments will be briefly discussed concerning their radiation susceptibility, but first some general observations are in order. Any instrument that depends upon particle ionization for its operation is particularly susceptible to RTG radiation. This includes scintillation devices, surface barrier solid-state detectors, and ionization chambers. In practice, most of the instruments for the scientific experiments do not depend on particle ionization or emission for proper functioning. Equally important is the fact that many of the experiments concern themselves with electromagnetic radiation at wavelengths quite different from those caused by an RTG, or are designed to measure phenomena completely unrelated to the RTG-produced fields. Examples of the

different-wavelength experiments include microwave, visual and ultraviolet radiometers and spectrometers, radio propagation and occultation experiments, and television experiments. Examples of experiments independent of RTG-caused fields are micrometeoroid detectors or cosmic dust collectors.

With these points in mind, the following observations concerning specific instrument types may be made:

a. Charged particle detectors. These include electrostatic and electromagnetic devices as well as ionization devices. The first two types, including Langmuir-type probes, RF impedance probes, Faraday cups, and electromagnetic analyzers, are inherently "hard" to RTG radiation, provided they are properly oriented with respect to the RTG. They are, on the other hand, susceptible to spacecraft charge effects. However, they are not suitable for many applications, and ionization-type detectors must be utilized. The surface-barrier gold-silicon detectors are superior to large volume scintillation detectors because of the reduced probability of a neutron or gamma ray interaction with small detector volumes. In general, instruments specifically designed to detect electrons employ smaller detector volumes and, therefore, are less susceptible to RTG effects.

b. Neutral particle detectors. These devices include mass spectrometers and hot- or cold-cathode ionization gauges. Instruments of these types are not appreciably affected by spurious RTG radiation, since the number of electrons necessarily generated by the instrument to induce ionization is far in excess of the number which might be generated by the secondary processes associated with the absorption of RTG radiations.

c. Photon detectors. No RTG-caused interference is expected with devices measuring electromagnetic radiation at optical or greater wavelengths. This includes radiometers and optical devices such as telescopes and television cameras. However, at shorter wavelengths RTG interference can be expected, especially with instruments designed to detect X-rays and gamma rays. This, of course, is due to the presence of these photons in the RTG radiation spectrum.

d. Magnetic field detectors. These devices include both flux gate magnetometers and the helium plasma magnetometer. Neither of these types of magnetometers is susceptible to RTG radiations, although stray magnetic fields which might result from the RTG must be reckoned with

if definitive magnetic field measurements are to be performed.

e. Photomultiplier tubes. Although not generally used as radiation detectors directly, photomultiplier tubes are an inherent part of scintillation counters and other devices which must respond to photon pulses in the optical region.

Thresholds have been established for gamma-ray-induced conductivity in some types of photomultiplier tubes. This induced conductivity is from two major sources: (1) luminescence induced in the glass envelope, and (2) increased electron density in the vicinity of the first few dynodes due to Compton scattering of the gamma rays. Testing has indicated that at least 5×10^5 gammas/cm²-s at 1 MeV are required to produce measurable conductivity. This flux of gamma rays is in excess of the number expected from an RTG at distances of more than 1 m.

4. Experiments

A listing of the experiments that will comprise the science package for a multiplanet mission is given in Table 11. Included in the table is an indication of whether the particular experiment will be affected by RTG radiation, and other relevant factors. Note that six of the experiments are insensitive to RTG radiation, and that two others are affected only under certain conditions. Among the latter two, the low-energy proton and electron differential energy analyzer (LEPEDEA) experiment will encounter interference only if this device is used to measure the small flux of electrons and protons that exists within the interplanetary environment. In general, this unit would be used to measure the charged particle environment surrounding the target planets. Under these conditions, the particle fluxes are several orders of magnitude above the RTG radiation levels.

The first six named experiments in Table 11, i.e., the micrometeorite detector, the television system, the infrared radiometer, the infrared interferometer, the solar plasma probe, and the magnetometer are not expected to have any interference from RTG radiation, and will be suitable for use without shielding or special consideration.

Experiment 7, the LEPEDEA, is described in Ref. 1. This instrument utilizes cylindrical curved-plate electrostatic analyzers to provide measurements of the differential energy spectrums of protons and electrons within and in the vicinity of the earth's magnetosphere. Continuous channel multipliers (Bendix "Channeltrons") are used to

Table 11. Susceptibility of science instruments to onboard RTG

Experiment	Possible effect	Nature of sensitive element	Remarks
1. Micrometeorite detector	No		
2. Television system	No		
3. Infrared radiometer	No		Needs simple radiant shielding
4. Infrared interferometer	No		Needs simple radiant shielding
5. Solar plasma probe	No		
6. Magnetometer	No		
7. LEPDEA	Maybe	Channeltron multiplier and Geiger-Muller tube	OK for particle environment in vicinity of planets; possible interference if used to measure the smaller fluxes in interplanetary environments
8. Ultraviolet photometer	Maybe	Channeltron multiplier and Geiger-Muller tube	Possible interference from secondary emission if entrance area unshielded
9. Trapped radiation detector	Yes	Solid-state detector and Geiger-Muller tube	
10. Cosmic ray and energetic solar charged particle experiment	Yes	Solid-state detector (Li-drift Si) (D_1 , D_2 , D_3)	RTG causes $500\times$ space rate for low energy (single channel)
		Plastic scintillation detector (D_4)	Small loss using coincidence
		Photomultiplier to read plastic scintillation detector	Sensitive photomultipliers may give false responses due to gamma-induced luminescence in glass envelope or false count from Compton scattered electrons in vicinity of initial dynode

count individual charged particles accepted by the analyzers and to provide the instrument with a dynamic range in proton and electron intensities extending from 10^4 to 10^{10} particles/cm²-s-sr in a given energy bandpass of the electrostatic analyzer. The widths of the energy bandpasses of the electrostatic analyzers are sufficiently wide to cover the entire energy range extending from 90 to 70,000 eV (protons and electrons separately) in 14 voltage steps on the curved plates. The four electrostatic analyzers (two analyzers each for protons and electrons covering the above energy range), complete with signal conditioner, high-voltage power supplies, and thermal shell, require an average power of 2 W and an instrumental weight of 6.3 lb.

Experiment 8 is the ultraviolet photometer. The wavelength regions of interest are the He I line at 584 Å, the He II line at 304 Å, the H α line at 1215 Å, and the OI line at 1305 Å. The secondary electrons ejected from a photocathode are amplified by passage through a channel multiplier. A combination of appropriate optical filters and surface preparation of the photocathodes will make each channel responsive only to the intended radiation, holding the dark current to under 1 count/s.

Experiment 9 is the trapped radiation detector. The radiation to be detected consists of the charged particles,

namely protons and electrons, that are trapped near a planet by that planet's magnetic field. If the sensors are sensitive enough and the RTG interference level is low enough, the detector can also measure the nontrapped, free flux of charged particles in interplanetary space. The detecting elements consist of the thin-window Geiger tubes to detect electrons with energies of 40 keV, and a solid-state surface barrier detector for the proton flux that will respond to protons with energies from 0.5 to 10 MeV.

Experiment 10 is the cosmic ray and energetic solar charged particle experiment. This experiment is designed to measure the flux and determine an energy spectrum for protons and alpha particles. Protons are detected and an energy spectrum presented for the range up to 180 MeV and, correspondingly, alpha particles from 2 to 360 MeV are recorded. The experiment has directional discrimination, so it can measure the flux magnitude, the energy distribution, and the direction of those primary cosmic rays, consisting of protons and alpha particles.

5. Radiation and Shielding

Gamma-ray and neutron spectra and intensities at the surface of the RTG are those due to Gingo (SPS 37-57, Vol. III, pp. 73-76). The RTG configuration considered in this study was assumed, as was the fuel age used, to be

1000 days. Attenuation coefficients for tungsten for thicknesses up to 1 in. in ¼-in. increments were also taken from SPS 37-57, Vol. III.

The field from the RTG can be reduced both by separation and by interposing metal sheets as shielding. For this purpose, tungsten sheet ($Z = 74$, density = 19 g/cm^3) or uranium sheet (natural uranium, depleted of U-235 and U-233, $Z = 92$, density = 18.5 g/cm^3) are most satisfactory. Computations in this article are based on the use of tungsten shielding, although additional effectiveness can be obtained by the use of uranium shielding, since the mass absorption coefficient is about 50% higher for the same mass/cm² in the range 0.2–0.6 MeV. The relative advantage decreases towards 3 MeV but is still appreciable.

As the neutron flux penetrates the metal shielding, it is also somewhat attenuated, but in the process of capturing and removing neutrons from the flux, additional gamma photons, called "capture gammas," with a characteristic spectral distribution are created. In practice, the total number of capture gammas is somewhat under 1% of the number of direct gammas, in the case of tungsten shielding up to an inch thick. If depleted uranium were to be used as shielding, the number of capture gammas would be approximately doubled, ranging up to 2% of the direct gammas, although of somewhat lower maximum energy.

6. Analytic Model of Sensitive Components

a. Solid-state detectors. Experiments 9 and 10, the trapped radiation detector and the cosmic ray and energetic solar charge particle experiment, use the effect of interaction of charged particles in the depleted zone of a solid-state detector. Gamma photons traversing the depletion zone do not register directly. However, gamma photons are registered indirectly and in an energy-dependent fashion, since photo-electrons, Compton scattered electrons, or pair-produced electrons or positrons are liberated when a gamma photon interacts with the silicon or germanium material in the detector. These liberated electrons may then register a count in the same manner that a desired event would otherwise do in such a system.

A study done by A. R. Jones (Ref. 2) gives experimental information on the counting rates of silicon junctions for gamma photons of 0.1 and 0.4 MeV energies. This enables normalization of the calculated relative response of silicon and germanium detectors, and the determination of the contribution of the different energy portions of the RTG

gamma fields to the backgrounds of the solid-state detectors when used in Experiments 9 and 10.

The data points given by Jones were based on the use of a silicon p-i-n junction of $100 \text{ mm}^2 \times 500 \text{ }\mu\text{m}$ deep. He observed 10^8 counts/R at 0.1 MeV and 10^7 counts/R at 0.4 MeV.

Using the absorption coefficients given by White (Ref. 3), the number of gammas/cm² that equal 1R was determined. Thus, we find that 10^8 counts/R at 0.1 MeV is 10^8 counts/ 2.36×10^{10} gamma photons, or 0.0042 counts/gamma photon. Similarly, 10^7 counts/R at 0.4 MeV is 10^7 counts/ 4.60×10^9 gamma photons, or 0.0022 counts/gamma photon at this energy.

The absorption probabilities for Ge and Si resulting from photoelectric, Compton scattering, and pair production contributions are shown in Table 12; these data were adapted from Goulding (Ref. 4). Of course, not every absorption event leads to a count, since the depth of the depletion region and the threshold voltage in the detector electronics affect the counting efficiency. In the case of Jones' measurement at 0.1 MeV, there were 0.0042 counts/gamma photon, although the absorption probability from Table 12 is 0.125. Thus, the possibility of recording a liberated electron is $0.0042/0.125 = 3.4\%$ at 0.1 MeV. Similarly, at 0.4 MeV there are 0.0022 counts/gamma photon, and the absorption probability from Table 12 is 0.064. The probability of rendering a liberated electron is $0.0022/0.064 = 3.4\%$ at 0.4 MeV, the same probability as at 0.1 MeV. The 3.4% probability, which can be expected to be constant for the complete gamma spectrum of interest, relates only to Jones' measurement on a 500- μm -thick depletion zone. For other thicknesses, the response will be proportional to the thickness of the depleted zone relative to the 500- μm thickness quoted. In practice, an experimental one-point determination is most desirable for determining background to be expected from a given gamma flux intensity of known spectral distribution.

Thus, it is possible to calculate the interference in a solid-state detector caused by a gamma flux of known energy spectrum. This can be done by multiplying the flux at each energy by the corresponding probability of interaction taken from Table 12, taking into account the thickness of the depleted zone, its area, and the probability of rendering a liberated electron within the depletion zone (assumed to be 3.4% for all energies).

Note that the probability of interaction falls rapidly with increasing gamma energy. The interactions of the

Table 12. Probability of absorption of γ -ray photons in silicon and germanium detectors (0.3-cm thick)

E, MeV	Silicon detector				Germanium detector			
	Photoelectric effect	Compton scattering	Pair production	Total	Photoelectric effect	Compton scattering	Pair production	Total
0.04	0.30	0.170	—	0.470	—	—	—	—
0.1	0.015	0.110	—	0.125	0.4200	0.280	—	0.700
0.25	0.001	0.075	—	0.076	0.0320	0.160	—	0.192
0.35	—	0.068	—	0.068	0.0130	0.150	—	0.163
0.45	—	0.060	—	0.060	0.0080	0.130	—	0.138
0.55	—	0.053	—	0.053	0.0450	0.120	—	0.124
0.65	—	0.050	—	0.050	0.0030	0.110	—	0.113
0.75	—	0.048	—	0.048	0.0022	0.105	—	0.107
0.85	—	0.046	—	0.046	0.0018	0.100	—	0.102
0.95	—	0.042	—	0.042	0.0014	0.098	—	0.099
1.1	—	0.040	—	0.040	0.0010	0.090	—	0.091
1.3	—	0.036	—	0.036	—	0.082	0.0010	0.083
1.5	—	0.033	—	0.033	—	0.080	0.0012	0.081
1.7	—	0.031	—	0.031	—	0.075	0.0017	0.077
1.9	—	0.030	—	0.030	—	0.070	0.0021	0.072
2.5	—	0.026	0.0010	0.027	—	0.060	0.0040	0.064
3.5	—	0.021	0.0014	0.022	—	0.052	0.0085	0.060
4.5	—	0.020	0.0020	0.022	—	0.047	0.0120	0.059
5.5	—	0.015	0.0030	0.018	—	0.042	0.0150	0.057
6.5	—	0.014	0.0040	0.018	—	0.040	0.0200	0.060

lowest energy gamma photons, which potentially give the greatest interference, are those of the gamma radiations that are most easily shielded by heavy metal shielding.

b. Geiger-Muller tubes. Experiments 7, 8, and 9, the low-energy proton and electron differential energy analyzer, the ultraviolet photometer, and the trapped radiation detector, use Geiger tubes in conjunction with Channeltron or with solid-state detectors. The response of the Geiger tube to gamma photons impinging on it is due almost entirely to those secondary electrons produced in the tube wall which penetrate into the sensitive volume of the tube.

Much experimental information is available concerning the response of Geiger tubes to gamma photons of various energies, as a function of cathode wall material. (See e.g., Bradt, et al. (Ref. 5) and Maier-Liebnitz (Ref. 6).)

The Geiger tube which will be used in the science experiments has a stainless steel wall, and potential interference with the experiments will be the effect of production of secondary electrons in this wall by gamma photons from the RTG. A theory developed by G. von Droste (Ref. 7) gives the number of secondary electrons traversing the Geiger counter as a result of impinging photons for aluminum ($Z = 13$), brass ($Z = 29$), zinc ($Z = 30$), or lead ($Z = 82$) cathodes. For the energy range of interest,

that is, for γ photons of energy over 0.3 MeV, curves relating to brass ($Z = 29$) may be used for calculations of stainless steel cathodes ($Z = 26$).

The contribution of gamma photons of under 0.3 MeV is not appreciable because any shielding that is used attenuates the low-energy photons very strongly and preferentially compared with those of energy over 0.3 MeV, and because the low-energy gammas that do reach the counter have a small effect compared to photons over 0.3 MeV.

Experimental curves for the number of secondary electrons produced in photon interaction with the walls of a Geiger counter, given by Bradt, et al. (Ref. 5), for brass cathodes, are as shown in Fig. 24, which also shows reasonable agreement with the theory of von Droste.

To calculate the interference in a Geiger counter caused by a gamma flux of known energy spectrum, the flux at each energy can be multiplied by the coefficient for secondary electron production in the wall materials. Data from Ref. 7 indicates that above 0.3 MeV stainless steel walls will have the same coefficient as brass walls. In Fig. 24, the experimental values for brass walls are available and cover most of the range of interest in the use of the RTG.

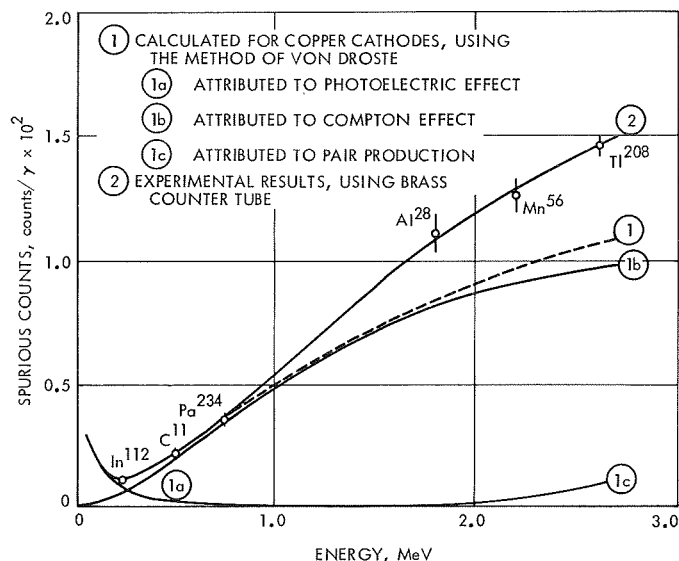


Fig. 24. Counting efficiency of Geiger counters with brass walls

Although the secondary electron emission coefficient rises rapidly above 3 MeV, the gamma spectrum of an RTG has few photons in the range 3–7 MeV, so errors in extrapolation of Fig. 24 do not impair the usefulness of the method.

c. Channeltrons and photomultipliers. Channeltrons are essentially distributed-dynode multipliers acting on the secondary electrons emitted by a target which is associated with but is not an integral part of the Channeltron. Photomultipliers are a combination of a (generally) low work function cathode with discrete dynodes which accomplish the electron amplification.

Both of these instruments depend on secondary electron emission for their normal operation, and are subject to increased background when the RTG-caused gamma flux causes additional secondary electrons to enter the dynode area. The material of the target, cathode, or other structure that interacts with the RTG-caused gamma flux is generally stainless steel, or stainless steel with such a thin coat of platinum or of alkali halides such as lithium fluoride, that the considerations developed for stainless steel cathodes in Geiger tubes hold here also, insofar as response versus gamma energy is concerned.

In the absence of specific experimental data, the coefficients given in Fig. 24 are useful to estimate the interaction of RTG gamma fields with these detectors.

7. Shielding Requirements and Effectiveness

The science experiment detectors, if located in the vicinity of an RTG, are subject to gamma and neutron fluxes which can give a background reading that will interfere with useful operation. The fluxes must be decreased to a tolerable value by the use of physical separation both from the RTG and from structural members of the spacecraft that can scatter gammas and neutrons, and by the use of shielding appropriately placed for maximum shielding effectiveness.

a. Channeltrons and Geiger counters. These detecting elements, which are used in the LEPDEA and in the ultraviolet photometer, respond to the background gamma field of the RTG by registering counts which originate from secondary electrons being ejected from the sensitive area of their inner surfaces.

As discussed earlier, the flux of gamma photons at each energy level can be multiplied by the coefficient for secondary electron production in the wall material to get the number of secondary electrons created per square centimeter of detector surface. If this is then multiplied by the sensitive area of the detector and by the efficiency η for Channeltron and Geiger counters, the number of background counts contributed by the RTG is obtained.

To calculate the background contributed by the RTG to the Channeltrons and to the Geiger counters, the following expression is evaluated:

$$R = \sum \eta \phi(E) \tau(E, x) \beta(E) f(d) A \quad (1)$$

where

R = spurious counting rate, counts/s

η = efficiency of counting ejected secondary electrons, which may be taken to be 100%

ϕ = photon flux at RTG surface, photons/cm²-s, as a function of energy

τ = transmission of the tungsten shield of thickness x , as a function of energy

β = secondary electron emission coefficient, as a function of energy (Fig. 24)

f = fraction of surface gamma intensity at a distance d from the RTG surface

A = sensitive area of detector, projected normal to flux

To arrive at the amount of shielding needed for Channeltron detectors, it should be noted that the sensitive

area of the Channeltron is the initial section of the tube. According to the manufacturer of the Channeltron 4010 tube, omission of the gain of the first 15% of the length of the tube decreases the output of the tube to 0.1 of design value, so it will suffice to shield this initial 15% of the 4-in. length, or 0.6-in. length. The inside diameter of the Channeltron 4010 is 0.040 in. To determine the effective sensitive area, we note that secondary electrons created by the passage of gamma rays through the walls of the Channeltron and through material surrounding the Channeltron can make their way through considerable thickness of wall material. The penetration of secondary electrons of 3 MeV is 2 g/cm², and this represents a wall thickness of 0.1 in. for material of the density of steel or brass, or correspondingly more for lower density wall material. Since 3-MeV secondary electrons represent the upper limit of the secondary electron spectrum that will be produced, the majority being considerably less penetrating, it is a conservative assumption to define the effective sensitive area of the Channeltron as a strip, 0.6 × 0.040 in., surrounded by a band 0.1 in. on all sides, or a total area of 0.8 × 0.24 in., which is 0.2 in.² or 1.2 cm².

When a Channeltron is used in the science experiment of a space vehicle, it is required to measure events when the background rate is 5 counts/min or 0.1 count/s.

Expression (1) has been evaluated for Channeltrons and is presented in Fig. 25. From this figure, one can determine the amount of tungsten shielding needed, for a range of separation distances, to reduce the background due to the gamma radiation from the RTG to the required level. If it is required to reduce the background count rate due to the RTG to be no greater than the background count rate due to all other causes, it must be reduced to 0.1 counts/s for the CEM-4010. This requires, as can be seen from Fig. 25, a separation of 10 ft if 2 in. of tungsten shielding is used. Other separations will require different thicknesses of tungsten, as shown.

The Geiger counter tubes that are to be used in the science package are specified as EON-6213, thin end window miniature self-quenching tubes. Although the secondary electron emission coefficient for these tubes with chromium-iron walls may be taken to be that given in Fig. 24, the difference in physical size of the sensitive area requires a separate evaluation of separation distance, shield thickness, and count rate interrelationships. The EON-6213 has an inside diameter of 0.093 in. and an effective length of 0.300 in. To determine the effective sensitive area, we note that the physical situation is simi-

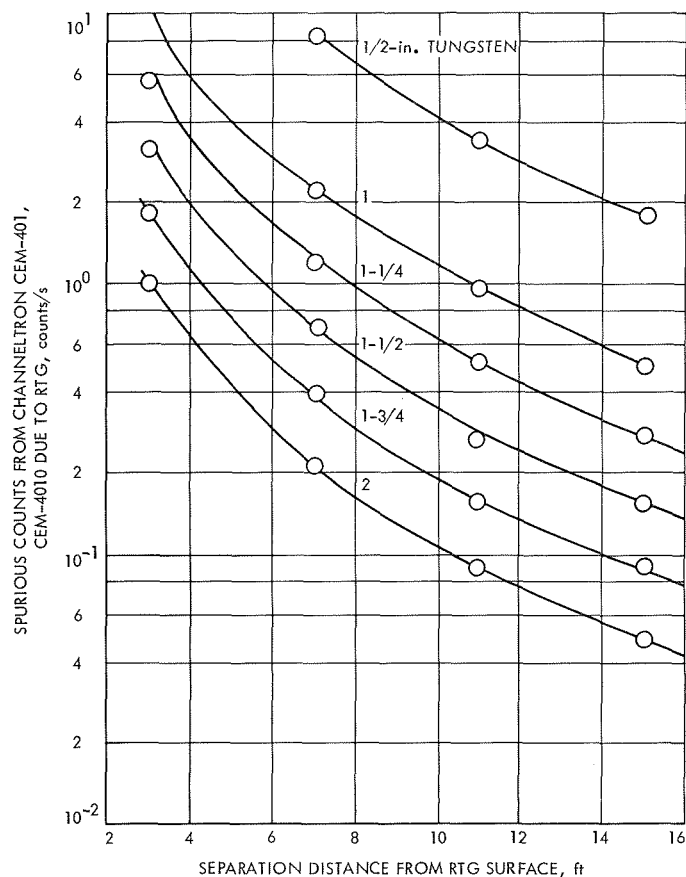


Fig. 25. Response of Channeltron CEM 401, CEM-4010 to RTG radiation

lar to that previously described for the Channeltron, wherein secondary electrons of 3-MeV energy can penetrate a wall thickness of 0.1 in. to appear in the inside gas volume of the Geiger tube. Thus, for the EON-6213 the conservative assumption is that the effective sensitive area is a strip, 0.093 × 0.300 in., surrounded by a band 0.1 in. on all sides, or a total area of 0.293 × 0.500 in., which is 0.15 in.² or 0.96 cm².

Expression (1) has been evaluated for these Geiger counters, and the results are presented in Fig. 26. As an example of the use of this figure, consider the trapped radiation detector, where the Geiger counters have a specified dynamic count range of 36 counts/min to 6 × 10⁸ counts/min. If the background due to the RTG is restricted to one-half the minimum dynamic range, i.e., 18 counts/min or 0.3 counts/s, it is seen that 1¼ in. of shielding at 13 ft or 1½ in. of shielding at 10 ft is required.

b. Solid-state detectors. Solid-state detectors will be used in the trapped radiation detector, and in the cosmic ray telescope. To calculate the background contributed

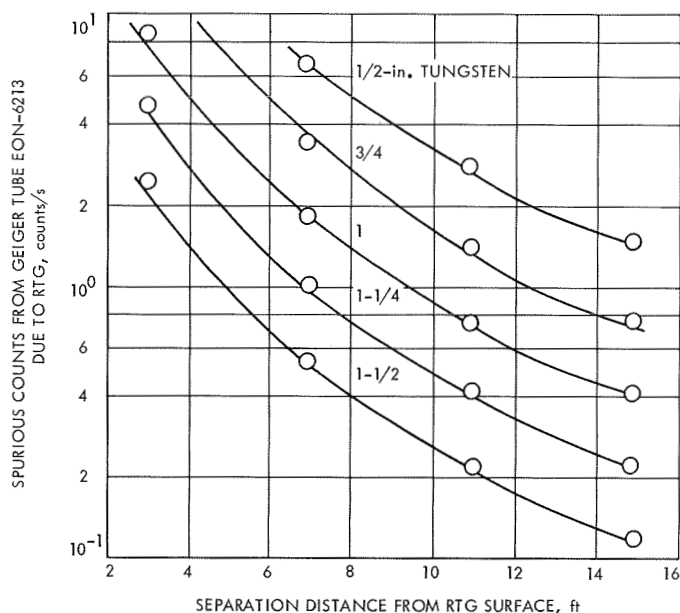


Fig. 26. Response of Geiger tube EON-6213 to RTG radiation

by the RTG to the solid-state detectors, use can be made of the experimental data given by Jones (Ref. 2) and reviewed earlier. It was shown that for a silicon detector of $100 \text{ mm}^2 \times 500 \text{ } \mu\text{m}$ depth (i.e., whose volume is $50,000 \text{ mm}^2 \times \mu\text{m}$ or 50 mm^3 or 0.05 cm^3), the number of counts produced was 3.4% of the number of photons absorbed in a 0.3-cm-thick slab of silicon. The 3.4% value held both at 0.1 and 0.4 MeV, and may be assumed to be valid for the entire spectrum of interest.

The product $\Sigma\phi_E\mu_E$, where ϕ_E is the gamma flux rate for photons of energy E and μ_E is the absorption coefficient for photons of energy E , was evaluated for the energy ranges of interest for a bare RTG and for shielding-modified fluxes. The values of μ_E were taken from Table 12. Then the number of counts to be expected from a $1 \text{ cm}^2 \times 500 \text{ } \mu\text{m}$ deep detector was calculated for each shield thickness.

Although the effect of shielding on flux-energy distribution was only carried out to 1 in. of tungsten in SPS 37-57, Vol. III, an extrapolation of the effectiveness of each additional $\frac{1}{4}$ -in. increment of tungsten shielding was made and allows these values to be carried out to the $2\frac{1}{2}$ in. of shielding needed in some cases. To arrive at the amount of shielding needed for the particular experiments, note that solid-state detectors are used in the trapped radiation detector experiment, and in the cosmic ray telescope experiment.

The information given on the trapped radiation detector has been summarized in Ref. 8. The silicon solid-state detector used has a normal sensitive area of 12 mm^2 and a depth of $31.7 \text{ } \mu\text{m}$. While the background or minimum needed rate is not quoted, the device has a "weak inflight source," providing 0.05 counts/s or less for calibration. If it is assumed that a background of 2% of the calibration strength is acceptable, shielding and distance must be adjusted to give 0.001 counts/s from the RTG. These calculations, based on Ref. 2, taking into account that the present silicon cell has a volume of $12 \text{ mm}^2 \times 31.7 \text{ } \mu\text{m}$, or only 0.0076 that of Jones' cell, and including the effect of separation distance, lead directly to Fig. 27, where background separation relationships are shown for several shield thicknesses.

The other experiment, the cosmic ray telescope, is described in Ref. 9, and the most recent embodiment is given in *Mariner Mars 1971 Quarterly Review No. 2*, Vol. 2, February 12-13, 1969. These indicate that the silicon detectors to be used will have a surface area of 5.7 cm^2

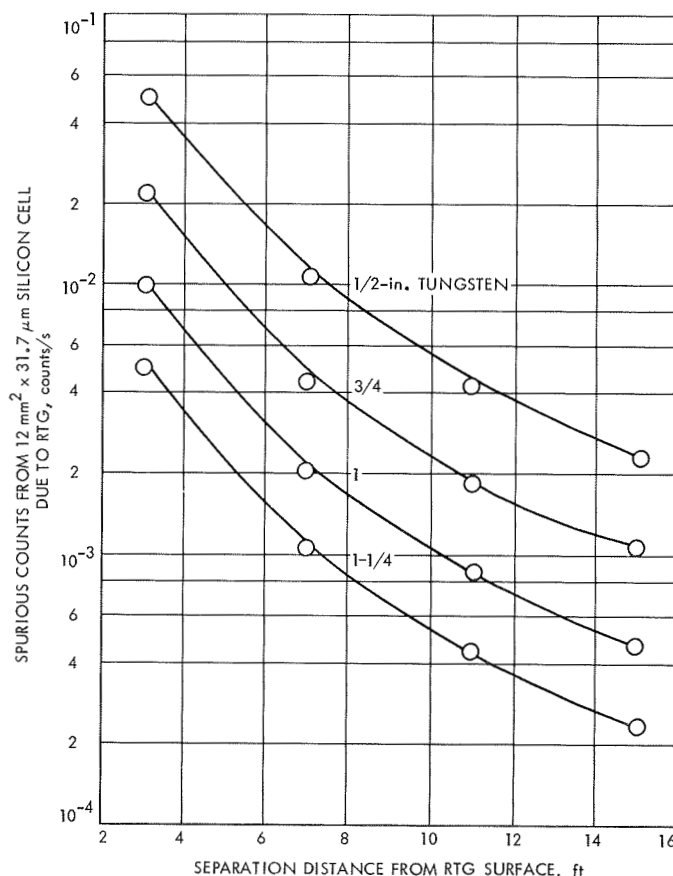


Fig. 27. Response of $12 \text{ mm}^2 \times 31.7 \text{ } \mu\text{m}$ silicon cell to RTG radiation

and a depletion depth of 200 μm . This value is 2.28 times that of Jones' cell, and the count rate to be expected must be multiplied by this factor for each shielding thickness.

A background count rate of 1 count/s will be under 10% of the minimum value of interest, and is suggested here as a tolerable RTG background. The background-versus-separation relationship is shown, for several thicknesses of shielding, in Fig. 28.

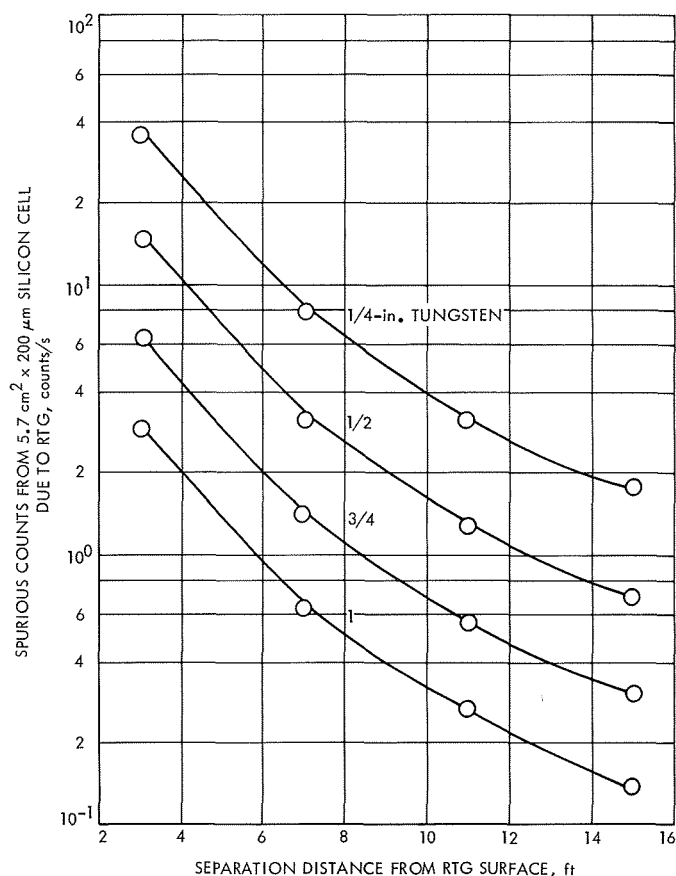


Fig. 28. Response of $5.7 \text{ cm}^2 \times 200 \mu\text{m}$ silicon cell to RTG radiation

c. *Shield mass for science experiments.* Table 13 gives a summary of required shield sizes and thicknesses to keep background rates of the individual experiments to the proposed limits. The practical size of each spot shield depends on the individual experiment and its container, taking into account accessibility of the components. Under some circumstances, a spot shield might be reduced in size, but such reduction depends on a detailed analysis of the container and enclosure for each experiment.

As was pointed out earlier, the use of depleted uranium as shielding in place of tungsten could be expected to reduce the required shield weight somewhat.

References

1. Frank, L. A., *Low-Energy Proton and Electron Experiment for the Orbiting Geophysical Observatories B and E*, Report 65-22. University of Iowa, Iowa City, Iowa, Jul. 1965.
2. Jones, A. R., *Gamma Ray Dosimetry with p-i-n Junction Converters*, AECL-2252. Atomic Energy Commission, Washington, D. C., Feb. 1965.
3. Gradstein, G. W., *X-ray Attenuation Coefficients from 10 keV to 100 MeV*, Circular 583. National Bureau of Standards, Washington, D. C., 1957.
4. Goulding, "Detectors in Radiation Energy Measurement," *IEEE Trans.*, NS 11, pp. 177-190, 1964.
5. Bradt, H., et al., "Empfindlichkeit von Zahlrohren mit Blei, Messing, und Aluminium-Kathode für γ -Strahlung in Energieintervall 0.1 Mev, bis 3 Mev," *Helv. Phys. Acta*, Vol. 19, p. 47, 1946.
6. Maier-Liebnitz, H., "Absolute Zahlrohrmessungen an γ -Strahlen," *Z. Naturforsch.*, Vol. 1, p. 243, 1946.
7. von Droste, G., *Z. Phys.*, Vol. 100, p. 529, 1936; Vol. 107, p. 474, 1936.
8. Canvel, H., *Trapped Radiation Detector*, Technical Report 32-1197. Jet Propulsion Laboratory, Pasadena, Calif., Nov. 15, 1967.
9. *Instruments and Spacecraft*, NASA SP-3028, p. 739. National Aeronautics and Space Administration, Washington, D. C., 1966.

Table 13. Shield weight required for use of science experiments in vicinity of RTG

Detector	LEPEDEA		Ultraviolet photometer		Trapped radiation detector		Cosmic ray telescope
	Channeltron (CEM-4010)	Geiger-Müller (EON-6213)	Channeltron (CEM-4010)	Geiger-Müller (EON-6213)	Solid-state (12 mm ² × 31 μm)	Geiger-Müller (EON-6213)	Solid state (5.7 cm ² × 200 μm)
Number required per experiment	2	1	1	1	1	3	1
Sensitive area to be protected, cm ²	1.2	1.0	1.2	1.0	0.12	1.0	5.7
Proposed spot shield size, cm ²	6	4	10	4	3	4	10
Background rate acceptable, counts/s	0.1	0.3	0.1	0.3	0.001	0.3	1
Shield thickness required at 11-ft separation:							
Tungsten, in.	2	1½	2	1½	1	1½	¾
Tungsten, in.	86.4	64.8	86.4	64.8	43.2	64.8	32.4
Shield thickness required at 15-ft separation:							
Tungsten, in.	1¾	1¼	1¾	1¼	¾	1¼	½
Tungsten, in.	75.6	54.0	75.6	54.0	32.4	54.0	21.6
Individual shield weights at 11-ft separation, g	518	259	864	259	130	259	324
^a Total shield weight at 11-ft separation, g	1036	259	864	259	130	777	324
Individual shield weights at 15-ft separation, g	454	216	756	216	97	216	216
^b Total shield weight at 15-ft separation, g	908	216	756	216	97	648	216
^a Total shield weight for 4 science experiments at 11-ft separation = 3649 g = 8.03 lb.							
^b Total shield weight for 4 science experiments at 15-ft separation = 3057 g = 6.73 lb.							

X. Guidance and Control Analysis and Integration

GUIDANCE AND CONTROL DIVISION

A. Orbit Insertion Maneuver Mechanization Study,

G. Jaivin

1. Introduction

Two different orbit insertion maneuver mechanizations are being studied. The first mechanization considered is the conventional fixed-attitude case in which the thrust vector remains pointed in a fixed-inertial direction during the engine burn. Such a maneuver is planned for the *Mariner* Mars 1971 mission. The second mechanization considered is the pitch-over maneuver, which consists of a commanded turn at a constant rate about the pitch axis during the thrust period.

The fixed-attitude case requires two commanded turns (roll and pitch) to position the engine thrust vector prior to the burn. The pitch-over case requires an additional roll turn to orient the spacecraft pitch axis perpendicular to the plane of motion (assuming an in-plane maneuver).

The pitch-over maneuver requires a smaller orbital velocity correction to achieve a given final orbit than the fixed-attitude case. The selectable pitch-over rate case,

while offering more versatility than a fixed pitch-over rate, is more complex to mechanize. Details of these findings are developed in the following subsections.

2. Fixed versus Selectable Pitch-Over Rates

Two alternate means of mechanizing the pitch-over during the burn are under investigation. The first method uses a fixed pitch-over rate built into the orbiter attitude-control system prior to launch. For each approach trajectory and Mars orbit, there is an optimum pitch-over rate that minimizes the orbit insertion velocity correction ΔV . Hence, in selecting the fixed rate to be used, the approach trajectory and Mars orbit must be known quite accurately at the time the hardware is built. The approach trajectory and Mars orbit, however, can be different if several spacecraft were launched and are most accurately known just prior to the time of the orbit insertion maneuver. Thus, although such a technique is relatively easy to implement, it has a disadvantage. Therefore, a second way of implementing the pitch-over is also being considered. This technique would permit the pitch-over rate to be arbitrarily selected at the time of the orbit-insertion maneuver. This would allow greater flexibility in selecting the Mars orbit

for each spacecraft, permit the building of identical spacecraft hardware and, possibly, permit a smaller ΔV allocation for orbit insertion.

3. Orbit Insertion ΔV

The minimum orbit insertion ΔV required has been determined as a function of apsidal rotation for the fixed-attitude and pitch-over maneuver cases. The assumptions and definitions used in the calculations are indicated in Table 1. The results obtained when a selectable pitch-over rate is assumed are shown in Fig. 1. The variation in the resultant pitch-over rate versus apsidal rotation for this case is indicated in Fig. 2. To assess the consequences of a fixed pitch-over rate, three rates were chosen, namely, 0.0203, 0.0176, and 0.0126 deg/s, which correspond to minimum ΔV maneuvers for apsidal rotations of 0, 12.5, and 25 deg, respectively. For each pitch-over rate the minimum orbit insertion ΔV maneuver was computed for several apsidal rotations. The results, which are given in Table 2, show that relatively small differences result be-

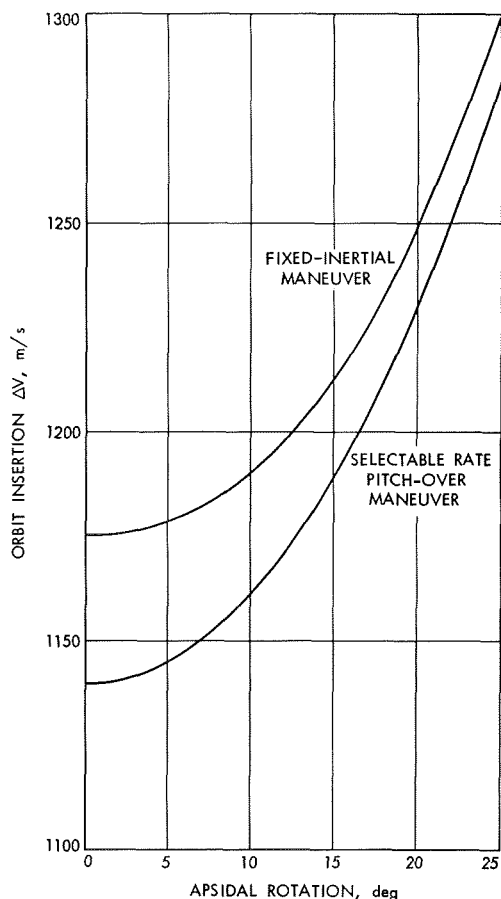


Fig. 1. Minimum orbit insertion ΔV for fixed-inertial and pitch-over maneuvers

tween the selectable and fixed pitch-over rates at the minimum ΔV point.

4. Implications of Trajectory Errors

If it were possible to always operate at the minimum ΔV point for the required apsidal rotation, the fixed pitch-over

Table 1. Assumptions and definitions

Item	Assumption or definition
Approach trajectory	Hyperbolic trajectory Variable impact parameter Hyperbolic excess velocity = 2.85 km/s
Mars orbit	Elliptic orbit Periapsis altitude = 1000 km Orbital period = 24.623 h
Maneuver	In-plane Engine thrust = 300 lb Specific impulse = 288 s Finite burn time
Spacecraft	Premaneuver weight = 7000 lb
Mars	Gravitational constant = 42828.44 km ³ /s ² Radius = 3393.4 km Spherical shape
Apsidal rotation	Apsidal rotation is defined as the angle between periapsis of a fiducial hyperbolic approach trajectory having both the actual hyperbolic excess velocity and a periapsis altitude of the final orbit and the line of apsides of the final elliptic orbit. The angle is measured positive in the direction of spacecraft motion. The angle between the asymptote of the actual approach hyperbola and the line of apsides of the final ellipse equals 123.058 deg + (apsidal rotation) for this case.

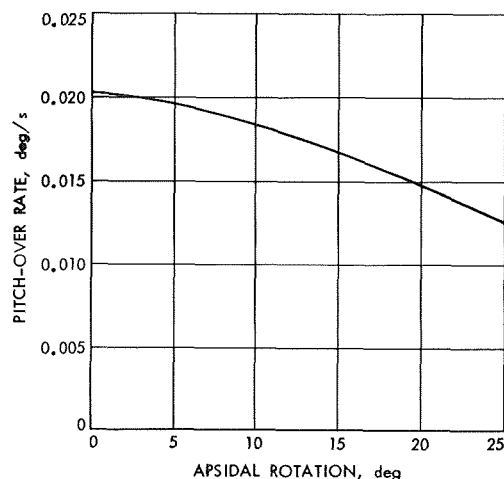


Fig. 2. Pitch-over rate variation for minimum orbit insertion ΔV case

Table 2. Minimum orbit insertion ΔV for several fixed pitch-over rates

Fixed pitch-over rate, deg/s	Apsidal rotation		
	0 deg	12.5 deg	25 deg
	ΔV , m/s		
0.0203 ^a	1140	1175	1289
0.0176 ^b	1140	1174	1285
0.0126 ^c	1146	1176	1282

^aOptimum rate for an apsidal rotation of 0 deg.
^bOptimum rate for an apsidal rotation of 12.5 deg.
^cOptimum rate for an apsidal rotation of 25 deg.

rate scheme would clearly be more desirable from the standpoint of simplicity of design. However, if consideration is given to errors in the approach trajectory, the situation changes. For instance, suppose a spacecraft were built with a fixed pitch-over rate of 0.0176 deg/s to optimize the orbit insertion maneuver for an apsidal rotation of 12.5 deg. As summarized in Table 3, the minimum ΔV required would be 1174 m/s. If, because of errors in the trajectory at the time of the orbit insertion maneuver, the approach trajectory impact parameter was 75 km less than desired, the minimum orbit insertion ΔV would be 1215 m/s. This is a 41 m/s increase over the optimum case. In the same situation, with a selectable pitch-over rate, the minimum ΔV would be 1178 m/s which is a change of only 4 m/s over the optimum case. The fixed-attitude maneuver minimum ΔV would be 1232 m/s for this case which is 32 m/s more than the optimum fixed-attitude maneuver. Thus, the selectable pitch-over rate mechanization is less sensitive to off-nominal conditions than either the fixed pitch-over rate case or the fixed-attitude case for this example. This result is true in general.

Moreover, as seen in Fig. 3 for this example, if the impact parameter were 75 km more than the nominal it would not be possible, with a fixed pitch-over rate of 0.0176 deg/s, to achieve a final orbit having the desired

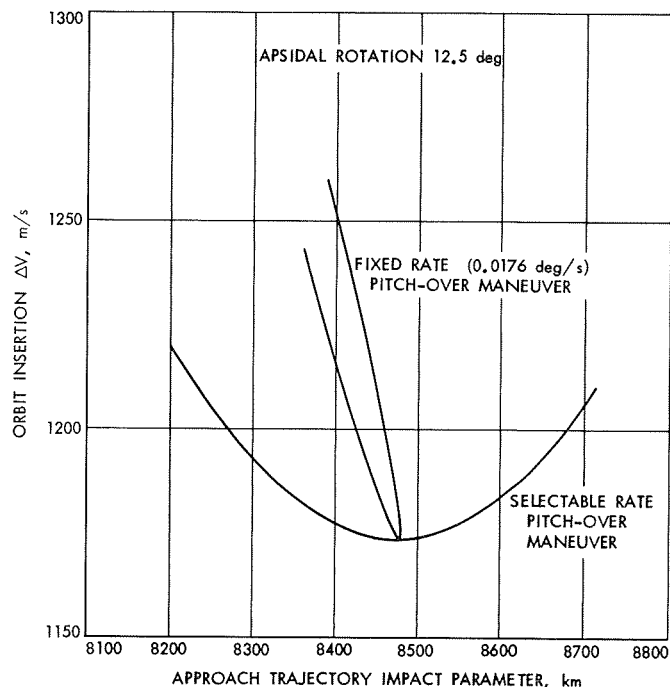


Fig. 3. Orbit insertion ΔV for fixed and selectable rate pitch-over maneuvers

apsidal rotation of 12.5 deg. This is also true of the fixed-attitude case. Therefore, it would be necessary, for such a mission, to bias the impact parameter aim point lower to assure the achievement of the desired apsidal rotation. Thus, a higher ΔV allotment would be needed for either the fixed-attitude case or the fixed pitch-over rate case than for the selectable rate case.

5. Variations in Impact Parameter

The approach trajectory impact parameters required for a minimum ΔV orbit insertion are larger for a pitch-over maneuver than for a fixed-attitude maneuver as indicated in Table 4. A larger impact parameter may be desirable from a planetary quarantine standpoint because the approach trajectory will not come as close to Mars.

Table 3. Typical example of trajectory errors for apsidal rotation = 12.5 deg

Error in impact parameter, km	Fixed-attitude maneuver, m/s	Fixed pitch-over rate maneuver (0.0176 deg/s), m/s	Selectable pitch-over rate maneuver, m/s
-75	1232	1215	1178
0	1200	1174	1174
75	"	"	1178

^aManeuver is not achievable.

Table 4. Impact parameters for minimum ΔV

Apsidal rotation, deg	Fixed-attitude maneuver, km	Pitch-over maneuver, km	
		Fixed rate ^a	Selectable rate
0	8060	8252	8280
12.5	8271	8477	8477
25	8941	9172	9110

^aPitch-over rate of 0.0176 deg/s (optimum for apsidal rotation of 12.5 deg). Other pitch-over rates show similar trends.

6. Pointing Error Sensitivity

If errors are introduced into both the initial thrust direction and the pitch-over rate, errors in the final orbit will result. Three orbit parameters, apsidal rotation, orbital period, and periapsis altitude, were examined with respect to such errors. The results, which are shown in Table 5, indicate that orbits with large apsidal rotations are most sensitive to such errors. The sensitivities of the orbital parameters to errors in the initial thrust direction are of comparable magnitudes in both the fixed-attitude and the pitch-over cases.

7. Discussion of Results

The pitch-over maneuver is, in general, a more efficient orbit-transfer mechanization than the fixed-attitude scheme from the standpoint of minimizing propellant consumption. It results in a larger impact parameter for a given maneuver. The pitch-over maneuver is more complex than the fixed-attitude case, particularly in the case of the selectable rate. The pitch-over maneuver may make

Table 5. Sensitivity of orbital parameters to pointing errors

\widetilde{ROT} , deg	$\frac{\partial \widetilde{ROT}}{\partial PI1}$	$\frac{\partial \widetilde{ROT}}{\partial PI2}$	$\frac{\partial TAU}{\partial PI1}$	$\frac{\partial TAU}{\partial PI2}$	$\frac{\partial HPE}{\partial PI1}$	$\frac{\partial HPE}{\partial PI2}$
0	-0.66	-0.78	0.04	-0.06	1.92	-6.2
12.5	-0.65	-0.78	-0.27	-0.45	16.3	11.2
25	-0.60	-0.77	-0.60	-0.93	32.0	32.3
HPE = periapsis altitude of ellipse, km. PI1 = initial thrust direction, deg. PI2 = pitch-over rate, millideg/s. \widetilde{ROT} = apsidal rotation, deg. TAU = orbital period, h.						

transmission of engineering data to earth more difficult during the engine burn.

These conflicting factors make the choice between the two mechanizations difficult. Further work is continuing to help delineate the advantages and drawbacks of the two schemes.

XI. Spacecraft Control

GUIDANCE AND CONTROL DIVISION

A. Digital Sun Sensor, *L. F. Schmidt*

1. Introduction

The sun sensor requirements for the Thermoelectric Outer Planet Spacecraft require design concepts not developed for prior missions. These requirements and the resulting design approach are described in SPS 37-57, Vol. III, pp. 103-108.

The basic approach is to form a line image of the sun on a digital detector through means of a cylindrical lens as shown in Fig. 1. The sensor output provides the sun angle in terms of a gray-coded digital word. Each bit of this word is generated by a plus or minus output voltage from the detector. The signal processing logic stores the digital word in a storage/shift register where it is sampled by the attitude control logic.

A single-axis breadboard having a 6-deg field-of-view (FOV) has been constructed and is now under test. The output from this sensor is an 8-bit digital word.

2. Signal Processing Electronics

The detector functions similar to a potentiometer. If equal positive and negative voltages are applied to the two ends of a linear potentiometer, the sliding contact

potential with respect to ground is analogous to the signal output from a typical bit on the detector as the line image of the sun moves on the detector bit. On the detector, light from the line image illuminates the photo-resistive material between the signal electrode and the positive excitation electrode lowering the resistance. Assuming that the opposite segment, which lies between the signal and negative excitation electrodes, is dark with a resulting high resistance, then a positive output potential results that is proportional to the resistances of the light and dark segments. For example, a 10:1 ratio of resistance will produce a peak-to-peak signal that is 80% of the excitation differential.

At the lower illumination levels anticipated at the outermost planets, the impedance of the detector will be in the megohm range. Consequently, the first step in the signal processing is to use an amplifier in a voltage follower configuration to buffer the detector output. This stage has a gain of one and a non-inverted output.

For simplification, assume that the detector signal for each bit is at the positive excitation level for a logical 1 and at the negative excitation level for a logical 0. The next stage in signal processing makes these logic levels compatible with the transistor-transistor logic and defines a specific trigger point for the bit change. To

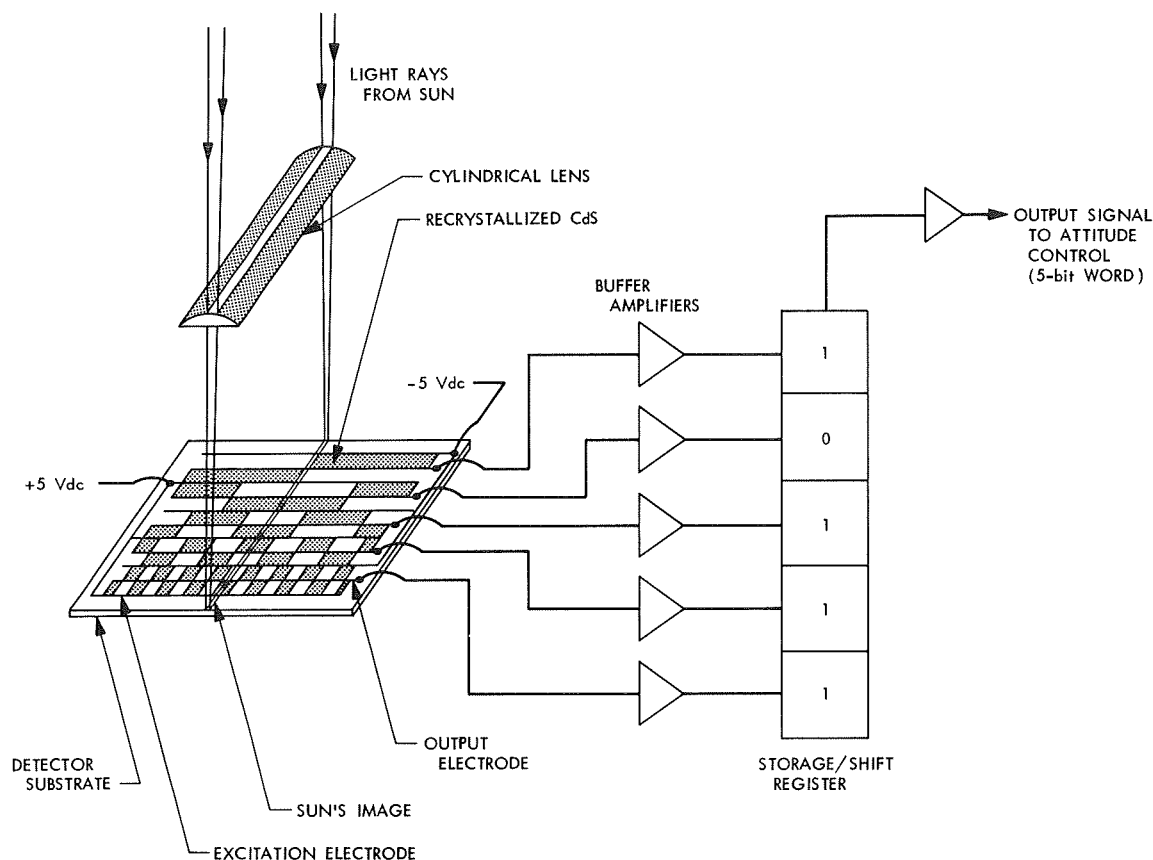


Fig. 1. Digital sun sensor conceptual diagram

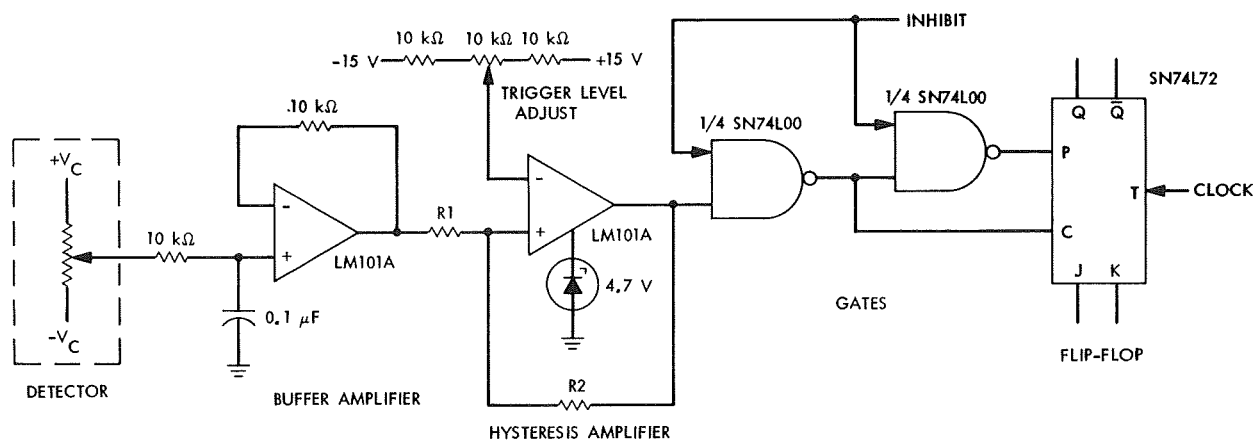


Fig. 2. Single bit signal processing circuit

accomplish this, an operational amplifier is used with positive feedback and with the output zener clamped at 0 and +4 V (Fig. 2). Hysteresis occurs as a function of the input resistor R_1 and feedback resistor R_2 . In determining the amount of hysteresis that is desirable, a tradeoff must be made between peak noise immunity at the bit transition and the shift of the bit edge when

the travel of the light beam changes direction. This shift is due to the finite slope of the transfer characteristic.

Following the hysteresis amplifier is a gated shift register for parallel-to-serial conversion during data transfer to the attitude-control electronics (ACE). There are two gates and a JK flip-flop for each bit as shown in

Fig. 2. The state of each bit (and its complement) is applied to the asynchronous inputs of the flip-flop, storing the bit states in real time. To transfer the data, the asynchronous inputs are inhibited and the shift register contents are shifted out, most significant digit first, into a storage register in ACE.

The ACE interface with the circuit in Fig. 2 will require a seven-wire cable: three excitation lines, a ground line, a clock line, an inhibit line, and a data line. Placing the shift registers for two axes in series eliminates the need for separate inhibit and clock lines for each axis. The clock and inhibit lines will be buffered and inverted at the sensor input due to fanout requirements and cable losses. The inversion allows the lines to remain at ground potential between data transfers. Likewise, the data line output is gated and buffered for isolation from the ACE.

The circuit shown in Fig. 2 has been incorporated into the single-axis breadboard and performs well. The trigger level adjustment shown was added to compensate for non-uniformities in bit response.

Figure 3 shows an alternate approach to the signal processing that has also been built and tested. It combines the hysteresis amplifier and inhibit gates in the Schmitt trigger; with the field-effect transistor, the operational amplifiers are eliminated entirely, reducing power consumption from 150 mW to less than 25 mW per bit.

The future development of the signal processing electronics will depend on design constraints imposed by environmental considerations, specifically neutron and gamma radiation, and the ultimate detector characteristics. Also, the availability of preferred microwatt logic

and operational amplifiers will affect the final electronic package configuration.

3. Breadboard Testing

a. Test method. Testing of the single-axis breadboard sensor to data has been conducted through use of a solar simulator for the stimulus. The sensor is located within the horizontal output beam from the simulator. Rotation of the sensor about a vertical axis has the effect of moving the sun through the sensor FOV. This test setup is shown in Fig. 4. The response of the detector is measured by recording the transfer function of each of the detector output signals that make up the 8-bit error-angle word.

Sensor accuracy is determined by rotating the sensor to a specific angle and comparing the gray-coded word from the signal processing logic to that angle. This is not practical for evaluation of the sensor over its entire FOV since the test would have to be repeated 256 times. It also would not accurately define the angles at which the 256-bit transitions occur. For tests of the prototype sensor, over 1500-bit transitions would have to be checked because of its larger FOV.

b. Test results. The design goal of 0.025-deg resolution from 6 to 30 AU has been achieved from 20 to 30 AU. The resolution from 1 to 20 AU varies as the sun's angular diameter changes and is approximately twice the design goal.

The width of the least significant bit (nominally 0.025 deg) varies over 0.025 deg under some conditions causing an improper sequencing of the output word. When this occurs, the resolution becomes 0.05 deg. The analysis of the detector transfer function indicates this problem is caused by non-uniformities in the detector characteristics.

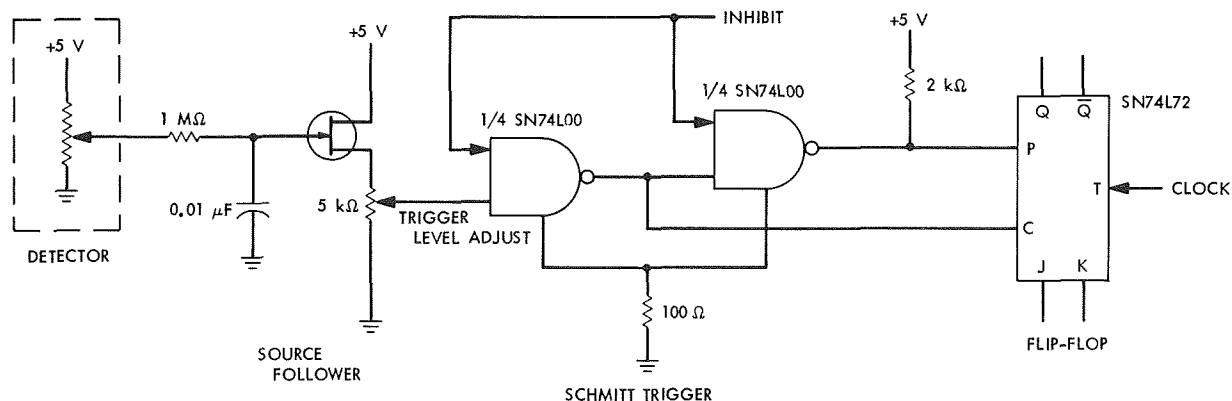


Fig. 3. Alternate signal processing circuit

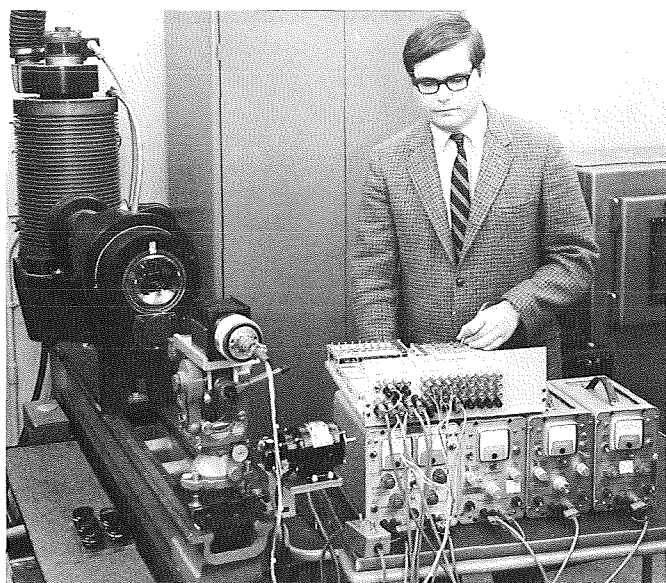


Fig. 4. Test setup

Improved detectors, which also have anti-reflection coated windows, are being incorporated into the bread-board and it is anticipated that improved resolution will be obtainable at higher illumination levels.

4. Solar Disc Simulator

Prototype sensor testing is scheduled on the sun simulator and at the celestial sensor high-altitude test facility located at Table Mountain. The Table Mountain facility contains a celeostat (servo-controlled mirror system) to produce a sunlit test area within a darkened room. For digital sun sensor testing, a solar disc simulator (SDS) is also required. The SDS operates in conjunction with the celeostat and simulates the sun's angular diameter from 1 to 30 AU. Figure 5 illustrates the configuration of the SDS. An aperture wheel is used to change the apparent angular diameter of the sun in eight discrete steps. It is estimated that the system will produce an illuminance in the test area of 0.8 solar constant com-

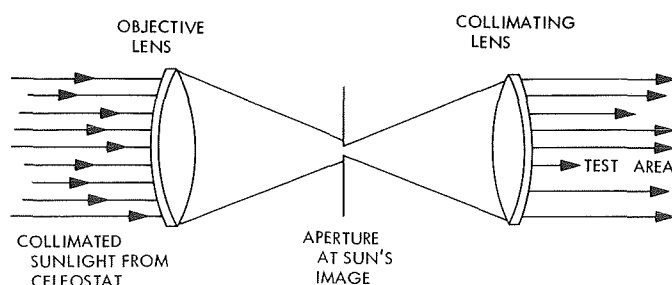


Fig. 5. Solar disc simulator

pared to a maximum of 0.5 solar constant for the sun simulator. The spectral mismatch of test area energy compared to the sun's energy as seen from space will also be much less of a problem when using the Table Mountain facility.

The optics for the SDS have been designed and are being fabricated. Installation at Table Mountain is scheduled for February 1970.

5. Data Acquisition and Processing System

The first objective in testing the prototype sensor is to determine its accuracy. A random comparison of the angle represented by the digital output word and the actual error angle results in a measure of accuracy that is limited by the finite width of the least significant bit. An improved test method would allow the determination of where the bit edges are in terms of actual error angle. When rotating the sensor such that the sun moves across its FOV, each change in output word indicates the occurrence of a bit transition. Since the accuracy of the sensor over the entire FOV is important, each of the approximately 1500-bit transitions must be accurately determined for each test run. An adequate number of test runs must be made to provide information for focusing of the optics, alignment of the detector with the optics, and alignment of the optical axis with the sensor reference surfaces.

Functional testing requires test and evaluation at various conditions of solar diameter and cross coupling. Repeat testing will be required after exposure to individual environmental conditions to determine the resulting degree of degradation, if any. An extensive amount of data must be acquired and reduced to test the sensor and to evaluate its performance. An automatic data acquisition system appears to be the most feasible solution to this problem of testing and evaluation.

The data acquisition system capable of accomplishing this task is comprised of several inter-related components. A motor-driven rate table, having a 17-bit binary encoder output, is used to rotate the sensor through its FOV. The axis of rotation is tilted to simulate cross-coupling conditions. The gray-code sensor output is converted to binary within a central digital processor. Each time the sensor output word changes, it is compared to the rate table encoder word. The angular deviation between the digital words is presented by a high-speed printer.

The ability of this system to acquire and process the test data in real time is extremely important. The first

advantage is the minimization of manpower required for testing. The second advantage is the minimization of test time necessary at Table Mountain. Since good seeing conditions are limited by weather conditions, it is important to accomplish as much as possible when favorable conditions occur.

B. Stability and Control of the Solar-Electric

Ion-Engine Propelled Spacecraft, L. L. Schumacher

1. Introduction

The solar-electric ion-engine propelled (SEP) spacecraft maintains 3-axis attitude control by thrust vector pointing control. The thrust vector is positioned relative to the spacecraft center of mass by translating the entire thruster array along two orthogonal axes in a plane perpendicular to the thrust vector, and by gimbaling the outermost ion thruster within the engine array.

A recent dynamic analysis of this configuration has revealed that translation of the engine array creates an attitude-control stability problem.

A command to translate the thruster array has two opposing effects: (1) a control torque is generated due to the engine thrust vector displaced from the spacecraft center of gravity, and (2) in translating the engine array, a torque is generated that tends to increase the spacecraft attitude error.

The sum of these torques determines the stability of the spacecraft. The attitude control system must ensure that the control torque is always largest, and that the difference between the two torques is large enough to ensure an adequate spacecraft response.

The destabilizing dynamic coupling effect was assumed negligible in the previous attitude-control design analysis. For this reason, a re-evaluation of the proposed attitude-control system was made. The re-evaluation was developed as follows:

- (1) The linearized rigid body SEP spacecraft dynamic equations were derived.
- (2) A comprehensive model of the coupled spacecraft and engine translator was developed. This model included dynamic contributions from the translator drive servo-motor that were neglected in the simplified dynamic analysis.
- (3) Trial methods of system compensation were selected for the comprehensive model.
- (4) The closed-loop characteristic equation was determined for each compensation scheme and it was then evaluated for the necessary and sufficient conditions for asymptotic stability.
- (5) The systems that satisfied the analytic stability criterion were then simulated on the UNIVAC 1108 computer using the CSSL dynamic simulation language. In addition, root locus plots were obtained so that possible system parameter improvements could be determined.

Results of each of these five steps are discussed in Subsections 2-5.

2. Development of the Linearized Rigid Body SEP Spacecraft Dynamic Equations

Figure 6 illustrates the model for the dynamic analysis. The motion is assumed to be in the plane of the paper and the spacecraft and the translator are lumped masses. The spacecraft mass M is assumed much larger than the translator mass m .

The X-Y coordinate system in Fig. 6 is assumed to be inertially fixed, and the coordinates of the spacecraft and translator centers of gravity are (X,Y) and (x,y) , respectively. The coordinate z indicates the position of the translator relative to some point on the spacecraft.

Forming the equations of motion using the methods of Lagrange, we arrive at the following equations of motion affecting spacecraft attitude:

$$J\ddot{\theta} + mc\ddot{z} = Fz \quad (1)$$

$$mc\ddot{\theta} + m\ddot{z} = B \quad (2)$$

where

$$J = (I + i + mc^2 + mz^2)$$

$$B = \text{force acting on the translator}$$

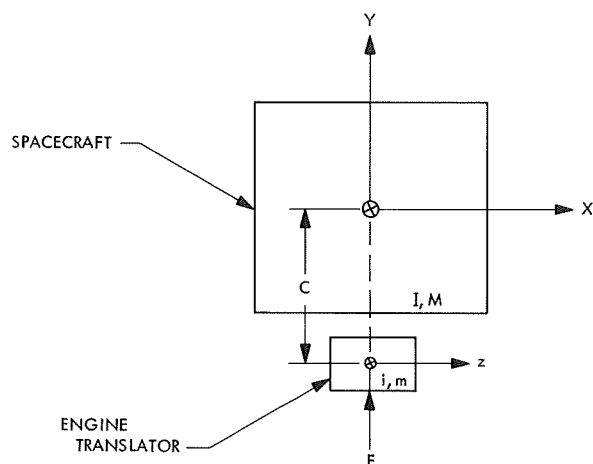
The parameters are defined in Fig. 6.

3. Development of a Comprehensive System Model

The comprehensive model developed contained the individual rigid body models of the spacecraft and thruster array translator with appropriate coupling forces.

The dynamic forces due to the interchange of momentum between spacecraft and thruster array mass are

(a) UNDISTURBED CASE



(b) DISTURBED CASE

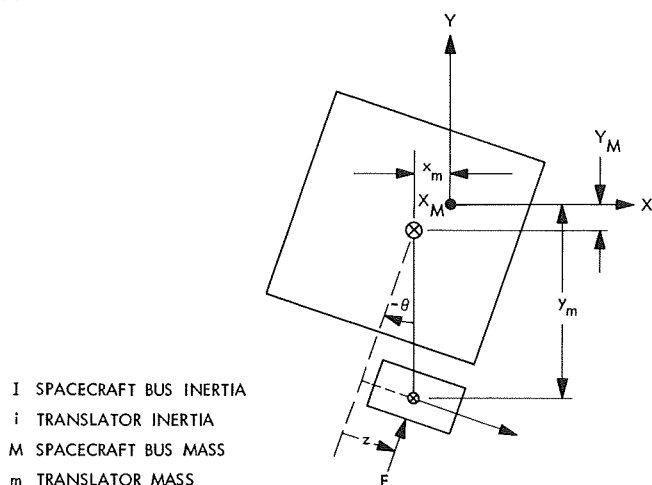


Fig. 6 Solar-electric spacecraft and engine translator rigid body model

explicitly stated in Eqs. (1) and (2). The forcing function (B of Eq. 2) contains the dynamics of the entire engine array positioning mechanism as well as position control commands. In order to determine the position control commands, it was first necessary to define the translator positioning mechanism dynamics.

The drive mechanism is basically a stepper motor and gear train. The detailed analysis of the stepper motor-thruster array mass dynamics revealed that the Laplace transform of $B(s)$ (Eq. 2) is as follows:

$$B(s) = C(s) - K_v s Z(s) - K_p Z(s) \quad (3)$$

where s is the Laplace operator. $Z(s)$ is the Laplace transform of the translator positions relative to a fixed point on the spacecraft. All friction and energy dissipation sources are lumped together as an equivalent linear viscous friction K_v , and the stepper-motor generated restoring forces are seen as linear stiffness term K_p . $C(s)$ is a function of the measureable spacecraft and/or translator variables and remains to be determined.

Figure 7 is a block diagram of the comprehensive model with no compensation added.

4. Selection of Trial Methods of Systems Compensation

In Eq. (3), $C(s)$ is the feedback control law. In this analysis, $C(s)$ will be a function of the continuously available sources spacecraft position, and/or thruster array position.

Figure 8 is the coupled spacecraft-translator block diagram and a feedback compensation scheme similar to that previously proposed. This diagram includes all devices necessary for successful hardware implementation and includes a voltage-controlled oscillator and a spacecraft position sensor filter amplifier.

The compensation scheme in Fig. 8 was one of the very few simple systems that could be shown to be asymptotically stable. The method of determining each system's stability is described in *Subsection 5*.

5. Stability Evaluation for the Rigid Body SEP Spacecraft Attitude-Control System

Many different feedback control law schemes were tried. The acceptability of each was determined in the following manner:

- (1) The closed-loop characteristic equation was evaluated, including the test feedback control law $C(s)$.
- (2) The characteristic equation was examined to determine if it could meet Routh's necessary condition for asymptotic stability.
- (3) If the necessary conditions were met, Routh's array was formed and, by parameter manipulation, attempts were made to satisfy Routh's sufficient condition for asymptotic stability.

The system proposed in Fig. 8 satisfies the analytic stability requirement for the properly selected parameter values.

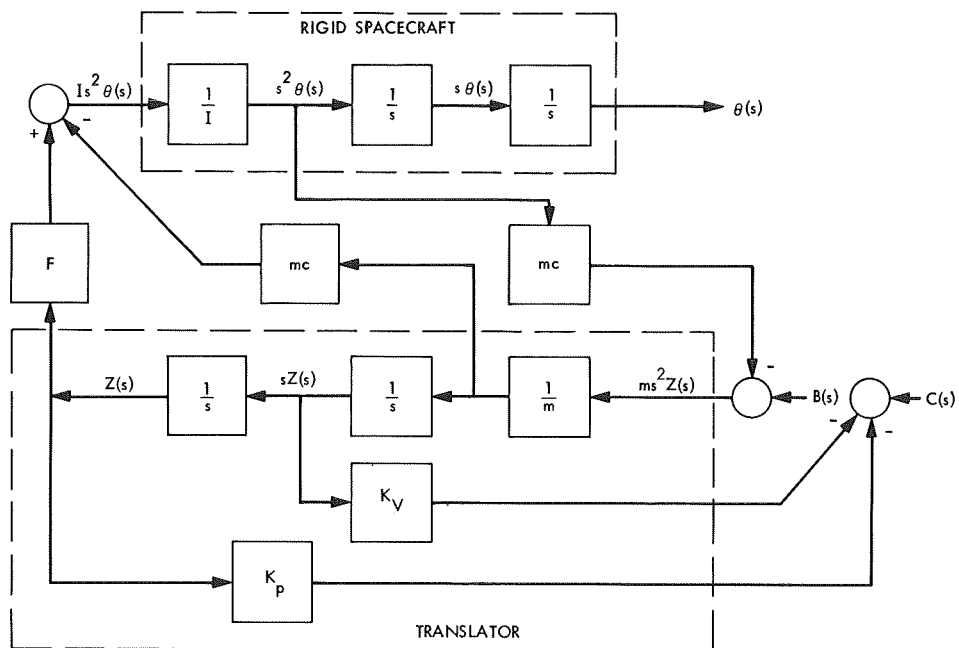


Fig. 7. Coupled spacecraft—translator model

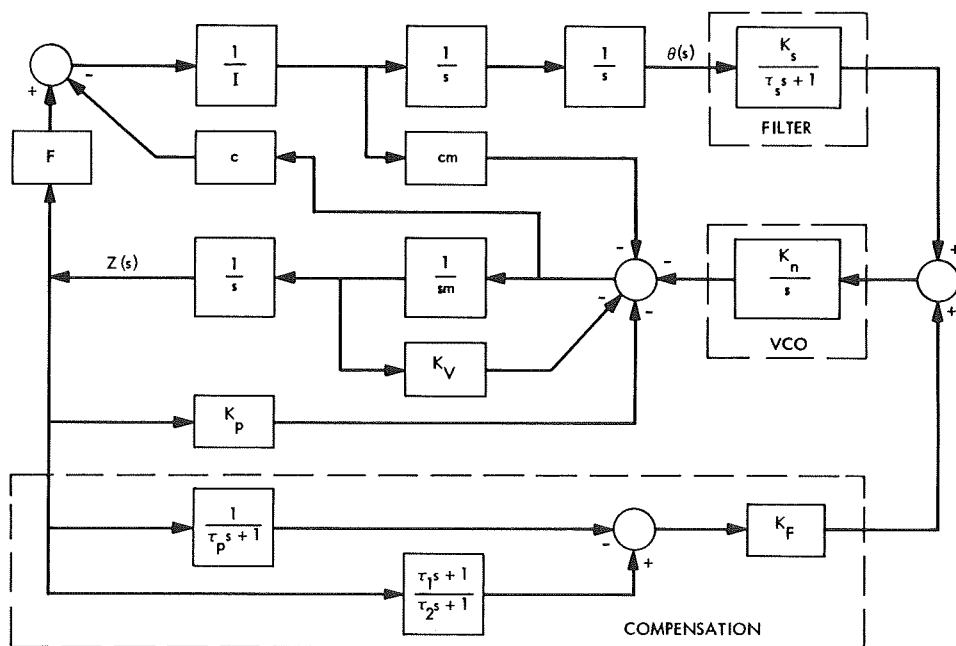


Fig. 8. Solar—electric spacecraft—translator with compensation

The analytic evaluation was verified with a dynamic simulation. Figure 9 is a plot of the phase plane for a disturbed rigid body model and the system shown in Fig. 8.

Figure 10 is a root locus plot of a large region of the s-plane and verifies the asymptotic stability prediction for a sufficient system gain.

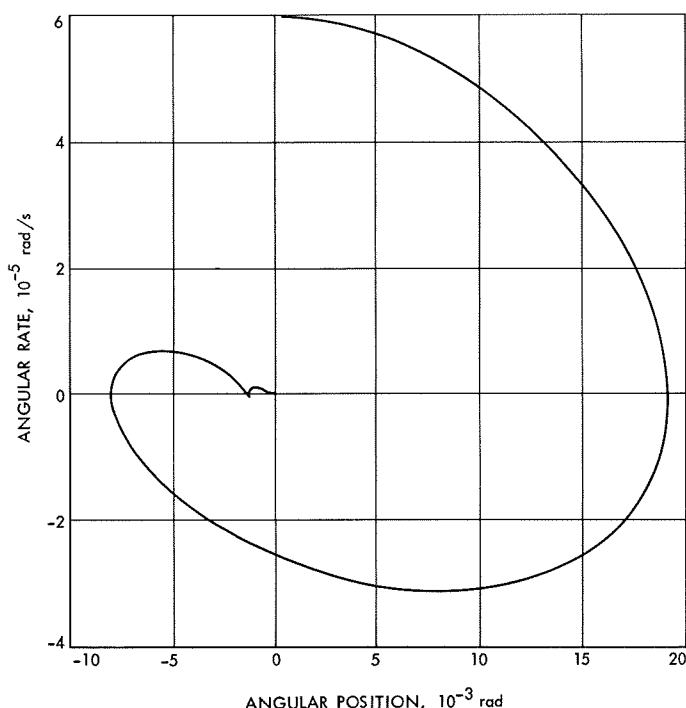


Fig. 9. Phase plans plot for rigid body model and translator with compensation

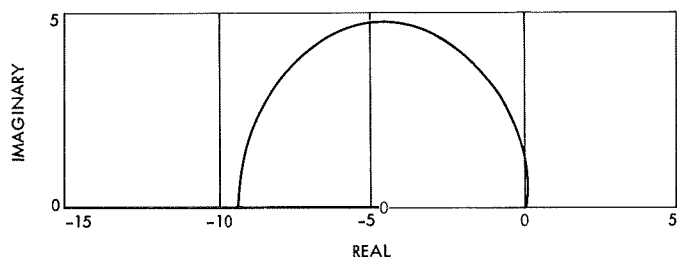


Fig. 10. Root locus plot in s-plane

6. Conclusion

The method described in this article is only one of the many methods of compensation that could be de-

veloped. It is believed that the comprehensive system model and techniques outlined herein will enable hardware designers to quickly determine other systems of compensation that may be more compatible with actual hardware implementation.

Areas needing additional investigation are as follows:

- (1) Investigate more compensation schemes.
- (2) Study interaction of flexible body dynamics with the compensated model.
- (3) Parameter study to optimize the damping characteristics so important in this type of spacecraft.

C. A Simple Digital-Derived-Rate Scheme, H. S. Lin

1. Introduction

It is well known in control theory that the optimal switching line of a second-order system is, in general, a nonlinear function of position and rate. Such a switching line can usually be provided with the help of a digital computer. Without the use of a computer, however, the generation of suitable rate information may be a difficult problem. In this article, a simple technique is developed that may be used to obtain the required rate information. This technique may be implemented in our present Thermoelectric Outer Planet Spacecraft attitude-control system discussed in SPS 37-59, Vol. III, pp. 148-149.

2. Position Error Signal

The position error is measured by a wide-angle digital sun sensor with an incremental resolution of 0.025 deg. This error signal is summed with a programmed central computer and sequencer (CC&S) signal that is used to bias the spacecraft's Z-axis away from the sun line to point the antenna to the earth. The resulting signal is sampled at 15-ms intervals and transmitted to a 12-bit storage register. The information is stored in a binary form (0.025 deg per increment) with the most significant bit (12th bit) indicating the polarity of the signal. It is this signal that is used to generate the necessary rate information.

3. Derived Rate

A derived-rate scheme has been used in all the *Mariner* missions to achieve the necessary passive stabilization during the cruise mode. However, the feedback signal generated by the *Mariner* derived-rate system is not proportional to the rate, but rather to the rate increment.

In other words, the term "derived rate" is a misnomer for that type of design. In order to obtain the proper rate information, the following capabilities must be provided:

- (1) The previous position error signal must be stored in the memory (flip-flop).
- (2) A change in the state of the position error must be detected.

Due to the high sampling rate (15 ms per sample), it is reasonable to assume that in all practical situations the position error will change by not more than one bit (0.025 deg) between any consecutive sampling intervals. Thus, only the two least-significant bits are required to derive the necessary rate information. The techniques used to obtain the direction and the magnitude of the rate information are discussed in the remainder of this subsection.

a. Direction of the derived rate. Figure 11 shows the direction decision logic. The signals PE1 through PE11 refer to the least-significant bit to the most-significant bit of the position error signal stored in the storage register. A change in error signal state is detected by a

differentiating circuit. When the state of the least-significant bit changes from zero to one or from one to zero, the corresponding logic signal ZTO or OTZ is generated by the differentiator. These signals are re-formed by a wave-shaping circuit and are *ored* together to produce the change-of-state signal PEC. The signal PEC will be used to transfer the current and previous position error data (PE2) into registers 1 and 2, respectively. The ZTODE and OTZDE signals are the logic signals ZTO and OTZ that have been passed through the delay (Fig. 11). The "true" and "false" outputs of registers 1 and 2 are processed by logic circuitry shown in Fig. 11 and are used to set or reset the direction indicator flip-flop. The function of the logic circuitry is straightforward. For example, if PE1 remains unchanged, a variation of PE1 from zero to one would indicate an increase in the position error, which corresponds to a positive angular rate.

b. Magnitude of the derived rate. Figure 12 shows the logic diagram used to determine the magnitude of the rate. A timing signal at a rate of approximately 1 pulse/s is generated from the 15-ms per pulse sampling signal. The time pulses are counted in an 8-bit counter

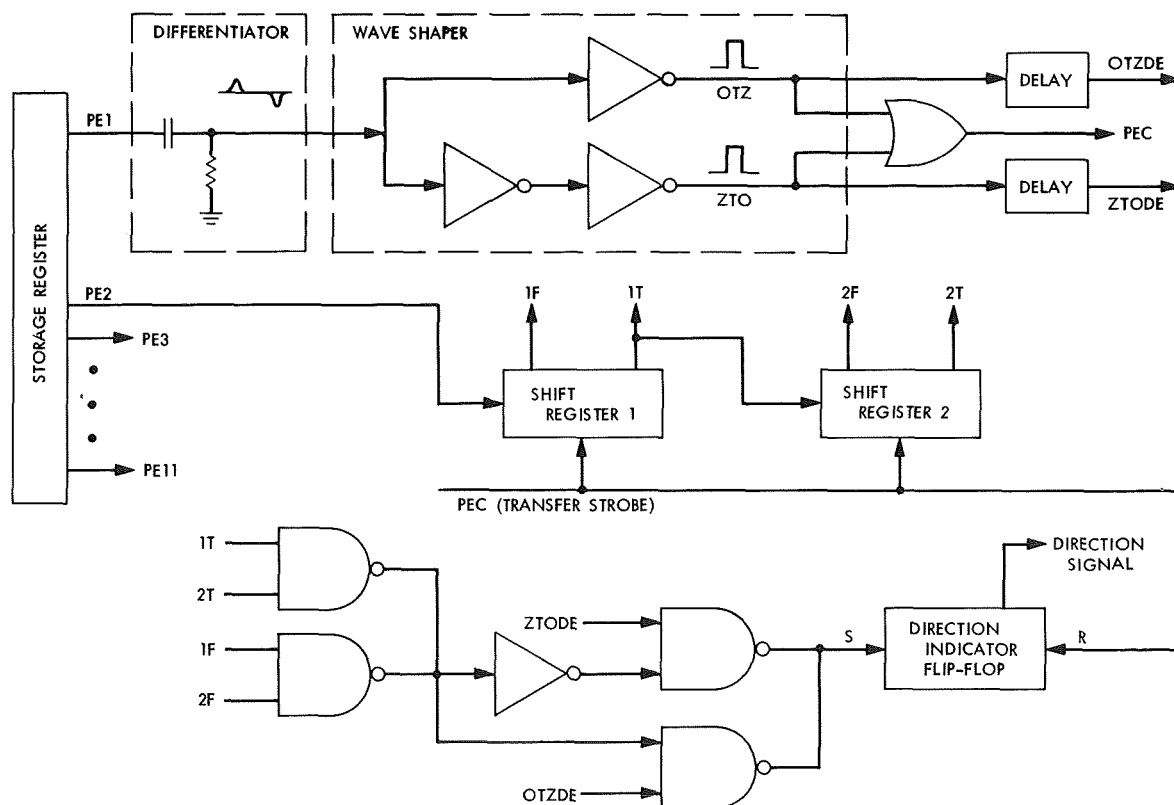


Fig. 11. Direction decision logic

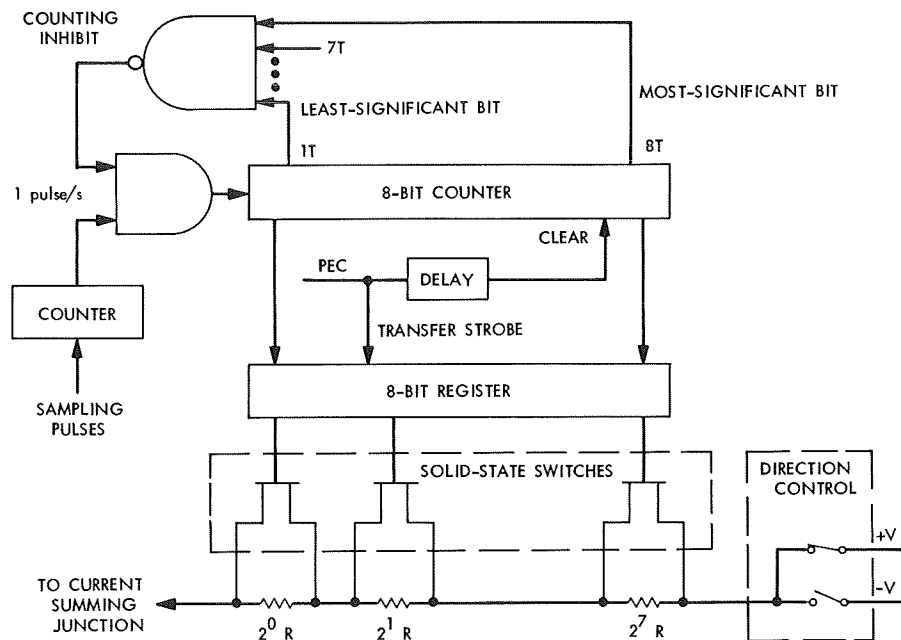


Fig. 12. Magnitude of the derived rate

from the moment when a change in position error state is detected. When the error state changes again, the information stored in the counter is transferred into an 8-bit register. After a short delay (less than 0.1 s), the counter is cleared and starts a new counting sequence. Thus, the state of the register indicates the time duration T between two consecutive position error state changes. This information is used to produce the necessary feedback signal. The outputs of the 8-bit register control eight normally closed solid-state switches that are in parallel with eight resistors. The impedance of the resistors varies in proportion to 2^i where i is the bit position. These switches will be opened if their corresponding registers are in a "one" state. Thus, the total resistance between the power supply and the feedback input summing junction is proportional to T . Therefore, the feedback current is proportional to the average angular rate of the spacecraft, which is the desired result.

4. Conclusion

In the previous subsection, a technique for deriving rate information has been developed. A similar approach can be applied to produce a signal that is proportional to the position error. With this information, a suboptimal linear switching line (Fig. 13) can be constructed. The selection of such a line should be based on the functional requirement of the system. This portion of the problem should be studied in the future.

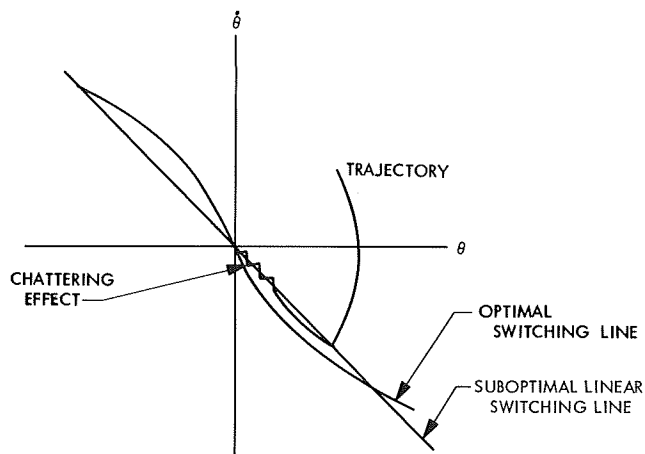


Fig. 13. Suboptimal linear switching

D. Nonlinear Control System Design Analysis via Nonlinear Filtering Techniques, A. K. Bejczy

1. Introduction

Application of optimal control (or, for that matter, of any control) requires a knowledge of the state variables of the dynamical system that has to be controlled. In practice, the system may be disturbed by random forces and the measurements may be contaminated by random errors. The basic question is then: How are the state variables inferred at a given time when both the system

and the measurements are subjected to random disturbances? From the point of view of control system design analysis, that basic question may be formulated as follows: For a given system and for a given set of realizable measurements related to the system, which measurements are necessary or preferable to obtain knowledge on the state variables within a prescribed limit of accuracy? Formulated alternatively, the basic question is: What improvement in the knowledge of the state variables can be accomplished by changing the configuration of the measurement vector and/or by increasing the number of components of the measurement vector? These questions are of great importance regarding complexity, cost, operating conditions and reliability of a control system.

The purpose of this article is to demonstrate that sequential filtering techniques may serve as effective analytical tools for investigating and/or selecting control system configurations from the point of view of the above considerations. Because of their practical significance, the discussion is confined to nonlinear systems. (Obviously, the main ideas carry through for linear systems, too.)

2. Performance Evaluation of Sequential Filters

Ideally, a sequential filter algorithm also describes the time history of the statistics of the filtering error in terms of known or knowable statistical parameters and of known system characteristics. Thus, evaluating the filters' performance for changing the configuration of the measurement vector and for varying the statistical parameters of the components of the measurement vector is essentially equivalent to a trade-off analysis for a control subsystem, namely for the measurement (or sensor) "box" of a control system.

3. Trade-Off Analysis for Nonlinear Systems

For *linear* systems, Kalman's filter algorithm provides a very compact, competent, and elegant performance evaluation method since it describes the complete statistics of the filtering error in a deterministic way. In the case of *nonlinear* systems, however, the known filter algorithms provide a probabilistic description of the (approximate) statistics of the filtering error. Evaluating the performance characteristics of nonlinear filters for varying statistical parameters and configuration of the measurement vector would normally require very extensive Monte Carlo type studies on nonlinear stochastic differential equations. In addition, the implementation of an entirely stochastic filter algorithm might not be feasible.

In SPS 37-59, Vol. III, pp. 132-137, a new performance evaluation method for approximate nonlinear filters was briefly outlined and illustrated on a scalar nonlinear problem. The new performance evaluation method provides a simple deterministic differential equation describing the (approximate) statistics of the nonlinear filtering error under certain assumptions. Moreover, the new method can also be applied to construct simple nonlinear filters that might yield as good response as the more complex nonlinear filter schemes. The new performance evaluation method, as it is outlined in SPS 37-59, Vol. III, will be applied in this article. It is also intended thereby to illustrate the salient features of the new method when it is applied to a *multivariable* nonlinear problem. The problem is selected from the domain of flexibly-coupled rigid bodies.

4. Nonlinear Spring System

Let the disturbed motion of a nonlinear damper with unit mass be described by the following system of ordinary nonlinear differential equations:

$$\dot{x}_1 = x_2 + n_{s_1} \quad (1)$$

$$\dot{x}_2 = -2x_1 - ax_1^3 - 3x_2 + 5\sin(t) + n_{s_2} \quad (2)$$

where the dot denotes time derivative, and

$$x_1 \triangleq \text{position}$$

$$x_2 \triangleq \text{velocity}$$

$$n_{s_1}, n_{s_2} \triangleq \text{random system noise}$$

The system noise is assumed to be uncorrelated Gaussian white noise with given spectral density (covariance matrix) denoted by

$$R = \begin{pmatrix} R_{11} & 0 \\ 0 & R_{22} \end{pmatrix} \quad (3)$$

Let the noisy measurements on position and velocity be given by a linear observation vector

$$y_1 = x_1 + n_{M_1} \quad (4)$$

$$y_2 = x_2 + n_{M_2} \quad (5)$$

where

$$n_{M_1}, n_{M_2} \triangleq \text{random measurement errors}$$

The measurement error is assumed to be uncorrelated Gaussian white noise with given spectral density (covariance matrix) denoted by

$$Q = \begin{pmatrix} Q_{11} & 0 \\ 0 & Q_{22} \end{pmatrix} \quad (6)$$

Furthermore, the measurement and system noise are assumed to be uncorrelated.

5. A Multidimensional Constant Gain Nonlinear Filter

For the approximate nonlinear filter, a constant filter gain matrix (denoted by G) will be postulated. According to Eq. (3) of SPS 37-59, Vol. III, the filter differential equations for sequentially estimating the mean of position and velocity of the nonlinear spring become

$$\dot{\hat{x}}_1 = \hat{x}_2 + G_{11}(y_1 - \hat{x}_1) + G_{12}(y_2 - \hat{x}_2) \quad (7)$$

$$\begin{aligned} \dot{\hat{x}}_2 = & -2\hat{x}_1 - a\hat{x}_1^3 - 3\hat{x}_2 + 5 \sin(t) \\ & + G_{21}(y_1 - \hat{x}_1) + G_{22}(y_2 - \hat{x}_2) \end{aligned} \quad (8)$$

with initial conditions

$$\hat{x}_1(0) = a_1$$

$$\hat{x}_2(0) = a_2$$

In subsequent calculations, for the constant a in Eq. (8), the value $a = 0.5$ will be used.

It is assumed that the mean of the filtering error μ is zero at $t = 0$. Then, according to Eq. (2) of SPS 37-59, Vol. III, the approximate covariance β of the filtering error is (approximately) described by the following system of ordinary linear differential equations:

$$\dot{\beta}_{11} = -2G_{11}\beta_{11} + 2(1 - G_{12})\beta_{12} + Q_{11}G_{11}^2 + Q_{22}G_{12}^2 + R_{11} \quad (9)$$

$$\begin{aligned} \dot{\beta}_{12} = & -(2 + G_{21})\beta_{11} - (3 + G_{11} + G_{22})\beta_{12} \\ & + (1 - G_{12})\beta_{22} + Q_{11}G_{11}G_{21} + Q_{22}G_{12}G_{22} \end{aligned} \quad (10)$$

$$\begin{aligned} \dot{\beta}_{22} = & -2(2 + G_{21})\beta_{12} - 2(3 + G_{22})\beta_{22} \\ & + Q_{11}G_{21}^2 + Q_{22}G_{22}^2 + R_{22} \end{aligned} \quad (11)$$

with given initial conditions

$$\beta_{11}(0) = b_{11}$$

$$\beta_{12}(0) = b_{12}$$

$$\beta_{22}(0) = b_{22}$$

Eqs. (9)–(11) establish a (deterministic) relation for the covariance of the filtering error β in terms of (1) filter gain G , (2) system noise R , and (3) measurement noise Q . The measurement noise covariance Q is a property of the sensing devices. Thus, Eqs. (9)–(11) also describe how the filter's performance (i.e., how the accuracy in the knowledge of the state variables) is affected by changing the accuracy properties and/or the functional configuration of the measuring devices. Consequently, Eqs. (9)–(11) may be viewed as an appropriate (though approximate) analytical tool for trade-off studies on the "sensor box" as a specified nonlinear control subsystem. The remaining question regarding the solution of Eqs. (9)–(11) is to select the constant G in some appropriate optimal sense.

The constant filter gains G_{ij} will be selected such that the steady-state value of the (positive definite) covariance matrix β is minimized. This requirement yields the following construction for the optimum constant gain matrix:

$$\begin{pmatrix} G_{11}^* & G_{12}^* \\ G_{21}^* & G_{22}^* \end{pmatrix} = \begin{pmatrix} \beta_{11}^* Q_{11}^{-1} & \beta_{12}^* Q_{22}^{-1} \\ \beta_{12}^* Q_{11}^{-1} & \beta_{22}^* Q_{22}^{-1} \end{pmatrix} \quad (12)$$

where the star denotes the minimum steady-state value of β_{ij} and the corresponding optimum constant value of G_{ij} , respectively.

The expressions for G_{ij}^* given by Eq. (12) substituted into the steady-state of Eqs. (9)–(11) yield the following quadratic (Riccati) system of algebraic equations for β_{ij}^* :

$$2\beta_{12}^* - Q_{11}^{-1}\beta_{11}^{*2} - Q_{22}^{-1}\beta_{12}^{*2} + R_{11} = 0 \quad (13)$$

$$-2\beta_{11}^* - 3\beta_{12}^* + \beta_{22}^* - Q_{11}^{-1}\beta_{11}^*\beta_{12}^* - Q_{22}^{-1}\beta_{12}^*\beta_{22}^* = 0 \quad (14)$$

$$-4\beta_{12}^* - 6\beta_{22}^* - Q_{11}^{-1}\beta_{12}^{*2} - Q_{22}^{-1}\beta_{22}^{*2} + R_{22} = 0 \quad (15)$$

Solving Eqs. (13)–(15)—by remembering the requirement of *minimum positive definite* value for β_{ij}^* —and substituting the resulting constant G_{ij}^* values given by

Eq. (12) into Eqs. (9)–(11), the proposed analytical tool for trade-off studies on the “sensor box” as a specified nonlinear control subsystem is complete in the sense of numerical application.

6. Numerical Results

In the subsequent application it is assumed that $R_{11} \equiv 0$. That is, there is system noise disturbing only the acceleration.

Obviously, there are three meaningful measurement configurations:

- (1) *Position measurements without velocity measurements.* In that case, terms containing Q_{22} drop out from Eqs. (13)–(15) as well as terms containing G_{12} , and G_{22} will be non-existent in Eqs. (7) and (8).
- (2) *Velocity measurements without position measurements.* In that case, terms containing Q_{11} drop out from Eqs. (13)–(15) as well as terms containing G_{11} , and G_{21} will be non-existent in Eqs. (7) and (8).
- (3) *Position and velocity measurements.* In that case Eqs. (13)–(15) and Eqs. (7) and (8) are fully applied.

Case 1: only Q_{11} . After some algebra, Eqs. (13)–(15) are reduced to solving the following fourth-order polynomial for β_{11}^* in the sense of finding its minimum positive real root:

$$12\beta_{11}^* + Q_{11}^{-1}\beta_{11}^{*2}(11 + 3Q_{11}^{-1}\beta_{11}^* + 0.25Q_{11}^{-2}\beta_{11}^{*2}) - R_{22} = 0 \quad (16)$$

Case 2: only Q_{22} . It is seen then from Eq. (13) that

$$\beta_{12}^* = 0 \quad (17)$$

Consequently,

$$\beta_{22}^* = 3Q_{22} \left[\left(1 + \frac{R_{22}}{9Q_{22}} \right)^{1/2} - 1 \right] \quad (18)$$

and

$$\beta_{11}^* = 0.5\beta_{22}^* \quad (19)$$

Case 3: Q_{11} and Q_{22} . In this case there are no simple algebraic expressions for β_{ij}^* . The simplest way of solving Eqs. (13)–(15) is by treating them as differential

equations. Clearly, they are then Riccati-type differential equations. Their equilibrium solution will be the solution of the algebraic system given by Eqs. (13)–(15). (This solution technique can also be applied in Cases 1 and 2.)

Table 1 contains some representative results regarding the performance of constant gain nonlinear filters for a case when the system noise has a fixed value ($R_{22} = 1.0$) and the configuration as well as the accuracy of the measurement vector are changed. Table 1 illustrates the significance of applying nonlinear filtering techniques for design analysis of nonlinear control systems. For instance, Table 1 reveals three interesting and important characteristics of the “measurement box” of a nonlinear control system in the case of a nonlinear spring: (1) Changing the accuracy of position and velocity measurements in the range of variance 0.5–4.0 (in appropriate units) would not affect significantly the accuracy of the knowledge on the state variables when constant gain nonlinear filters are applied. (2) Changing the configuration of the measurement vector (that is, having only either position or velocity measurements, or having both of them) would again not influence significantly the accuracy of the knowledge on the state variables when constant gain nonlinear filters are applied. In this regard, it is interesting to note that the simplest constant gain nonlinear filter structure can be obtained by

Table 1. The MSSVCG filter's performance for a nonlinear spring

	$Q_{11} = 0.5$ —	— $Q_{22} = 0.5$	$Q_{11} = 0.5$ $Q_{22} = 0.5$	$Q_{11} = 0.5$ $Q_{22} = 2.0$	$Q_{11} = 2.0$ $Q_{22} = 0.5$
β_{11}^*	0.073	0.079	0.069	0.072	0.076
β_{22}^*	0.163	0.158	0.155	0.161	0.157
	$Q_{11} = 1.0$ —	— $Q_{22} = 1.0$	$Q_{11} = 1.0$ $Q_{22} = 1.0$	$Q_{11} = 0.5$ $Q_{22} = 1.0$	$Q_{11} = 1.0$ $Q_{22} = 0.5$
β_{11}^*	0.078	0.081	0.076	0.071	0.074
β_{22}^*	0.165	0.162	0.160	0.159	0.157
	$Q_{11} = 4.0$ —	— $Q_{22} = 4.0$	$Q_{11} = 4.0$ $Q_{22} = 4.0$	$Q_{11} = 1.0$ $Q_{22} = 4.0$	$Q_{11} = 4.0$ $Q_{22} = 1.0$
β_{11}^*	0.082	0.083	0.081	0.077	0.080
β_{22}^*	0.166	0.166	0.165	0.164	0.162
NOTES: System noise is fixed, $R_{22} = 1.0$. R_{11} , Q_{ij} , and β_{ij}^* in appropriate units.					

measuring only velocity. In that case, $\beta_{12}^* = G_{12}^* = 0$ according to Eqs. (17) and (12), which obviously simplifies the implementation of the constant gain nonlinear filter. This means that, in that case, the best filtering result can be obtained if velocity measurements are *not* fed into the position-estimator channel of the filter. Regarding position estimates, we are better off by simply integrating over the estimated velocity when only velocity is measured. (3) However, from Table 1, it is clear that any constant gain nonlinear filter configuration will *substantially improve* the accuracy of the knowledge on the state variables as compared to the accuracy of the (noisy) measurements on the states.

Figures 14–17 illustrate some important and interesting details regarding the construction and performance

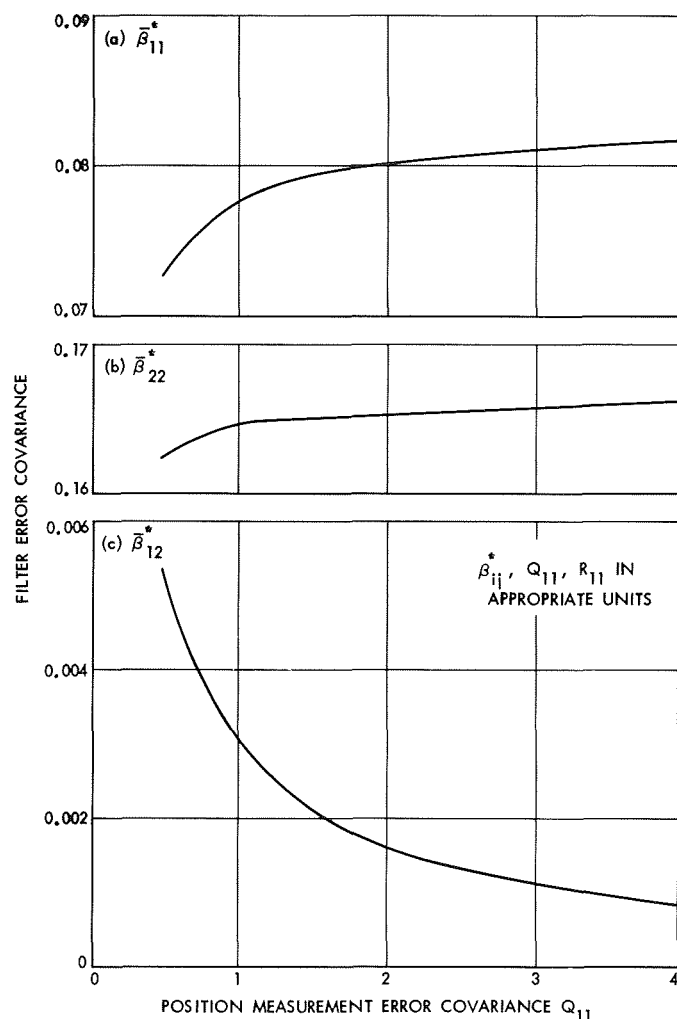


Fig. 14. Minimum steady-state covariance of a constant gain filter for a nonlinear spring (varying measurement noise, fixed system noise, $R_{22} = 1$)

of the minimum steady-state variance constant gain (MSSVCG) nonlinear filters as they are applied to the nonlinear spring problem when only (noisy) position measurements are used for estimation purposes.

Figure 14 depicts the filter's covariance β_{ij} as a function of the variance Q_{11} of position measurements. This

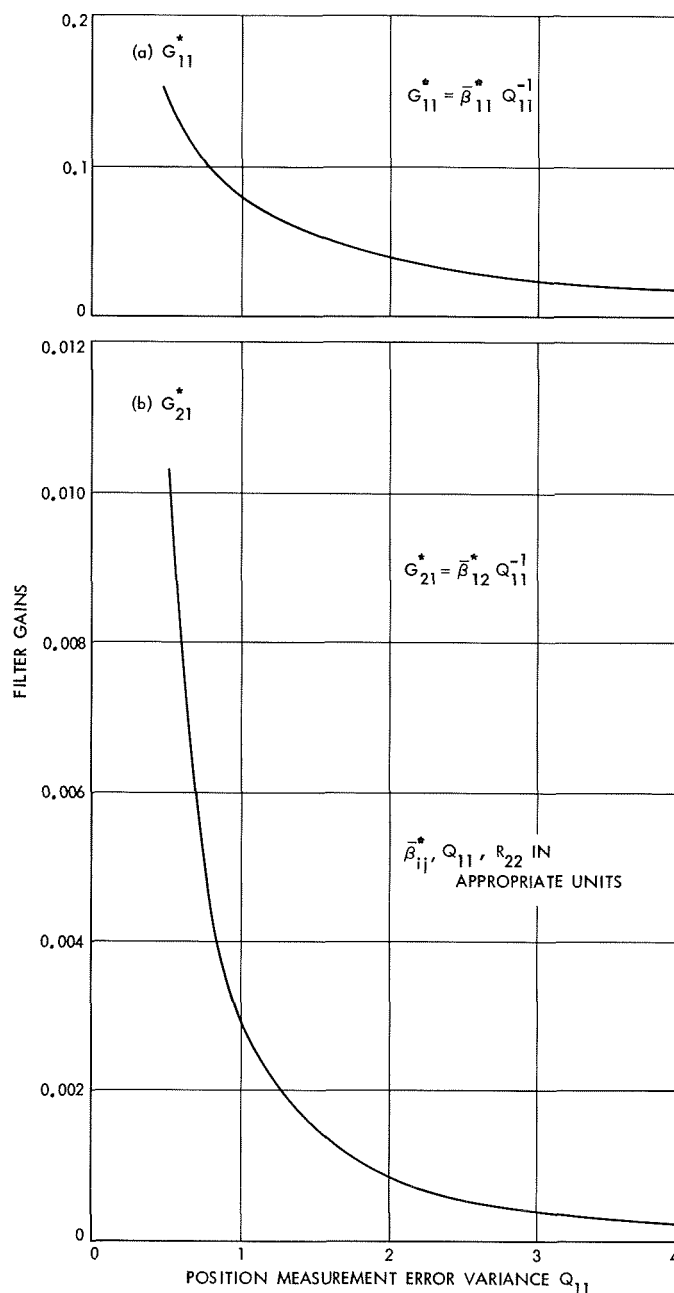


Fig. 15. Optimum constant filter gains for a nonlinear spring (varying measurement noise, fixed system noise, $R_{22} = 1$)

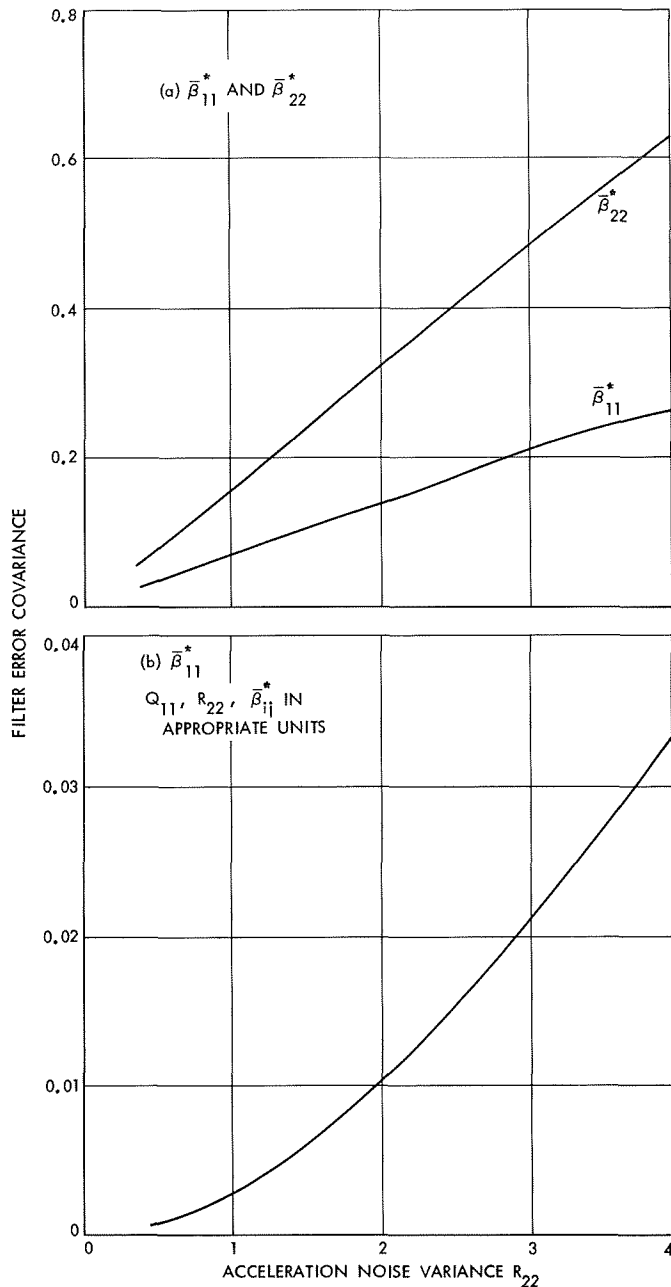


Fig. 16. Minimum steady-state covariance of a constant gain filter for a nonlinear spring (varying system noise, fixed measurement noise, $Q_{11} = 1$)

figure shows that the measurement noise can be filtered out very effectively by the MSSCG nonlinear filter. There will be no substantial change in the variance of the estimated position and velocity even though the variance of the position measurements is increased by a factor four.

Figure 15 shows the optimum constant gains G_{ij}^* as functions of the variance Q_{11} of position measurements.

Note that the G_{ij}^* values are changing (decreasing) roughly by a factor ten and twenty, respectively, when Q_{11} is increased by a factor four.

Figure 16 depicts the MSSVCG nonlinear filter's covariance β_{ij}^* as a function of the system noise (acceleration noise) variance R_{22} for a fixed variance of position measurements, $Q_{11} = 1.0$. Obviously, the filter's performance strongly depends upon the intensity of the system noise. The variance of the estimated position and velocity (β_{11}^* and β_{22}^* , respectively) are increasing roughly linearly with increasing R_{22} . It has to be emphasized that the *not faster* than linear increase in β_{11}^* and β_{22}^* with increasing R_{22} is a sign for the good quality or the MSSVCG nonlinear filter's structural strength. It is noted that, according to Eq. (12), the β_{11}^* and β_{12}^* functions depicted on Fig. 16 are, at the same time, numerically equal to the optimum constant gains G_{11}^* and G_{21}^* .

Figure 17 shows a sample case of the MSSVCG nonlinear filter as it is applied to the disturbed nonlinear spring problem having only (noisy) position measurements with error variance $Q_{11} = 1.0$ and acceleration noise with variance $R_{22} = 1.0$. That sample case is one of the 360 actual realizations of the filter by Monte Carlo simulation (MCS) on a digital computer. (The MCS for the filter is briefly described in SPS 37-59, Vol. III. The only change now is the increase of the number of variables.) Figure 17 depicts the "true" disturbed velocity and position trajectories and the "estimated" velocity and position trajectories together with the σ (standard deviation) channels belonging to the estimated position and velocity. The σ channels are computed from the appropriate solution of Eqs. (9)–(11). Thus,

$$\sigma(t)_{\text{position}} = [\beta_{11}^*(t)]^{1/2}$$

$$\sigma(t)_{\text{velocity}} = [\beta_{22}^*(t)]^{1/2}$$

Figure 17 also depicts the noisy position measurements and shows very impressively the performance characteristics of the MSSVCG nonlinear filter. Despite the strong measurement error (Fig. 17a) and despite the strong dynamical disturbance (Fig. 17b), the filter estimates the mean of position and velocity very effectively. It tracks the "true" trajectories after 1–2 s transient time. It also assures that reasonable (about 85–95%) portions of the "true" trajectories are inside the filter's own σ channels, which are centered around the estimated mean values. Note also the fine smoothing property of the MSSVCG nonlinear filter, which is clearly demonstrated in Fig. 17.

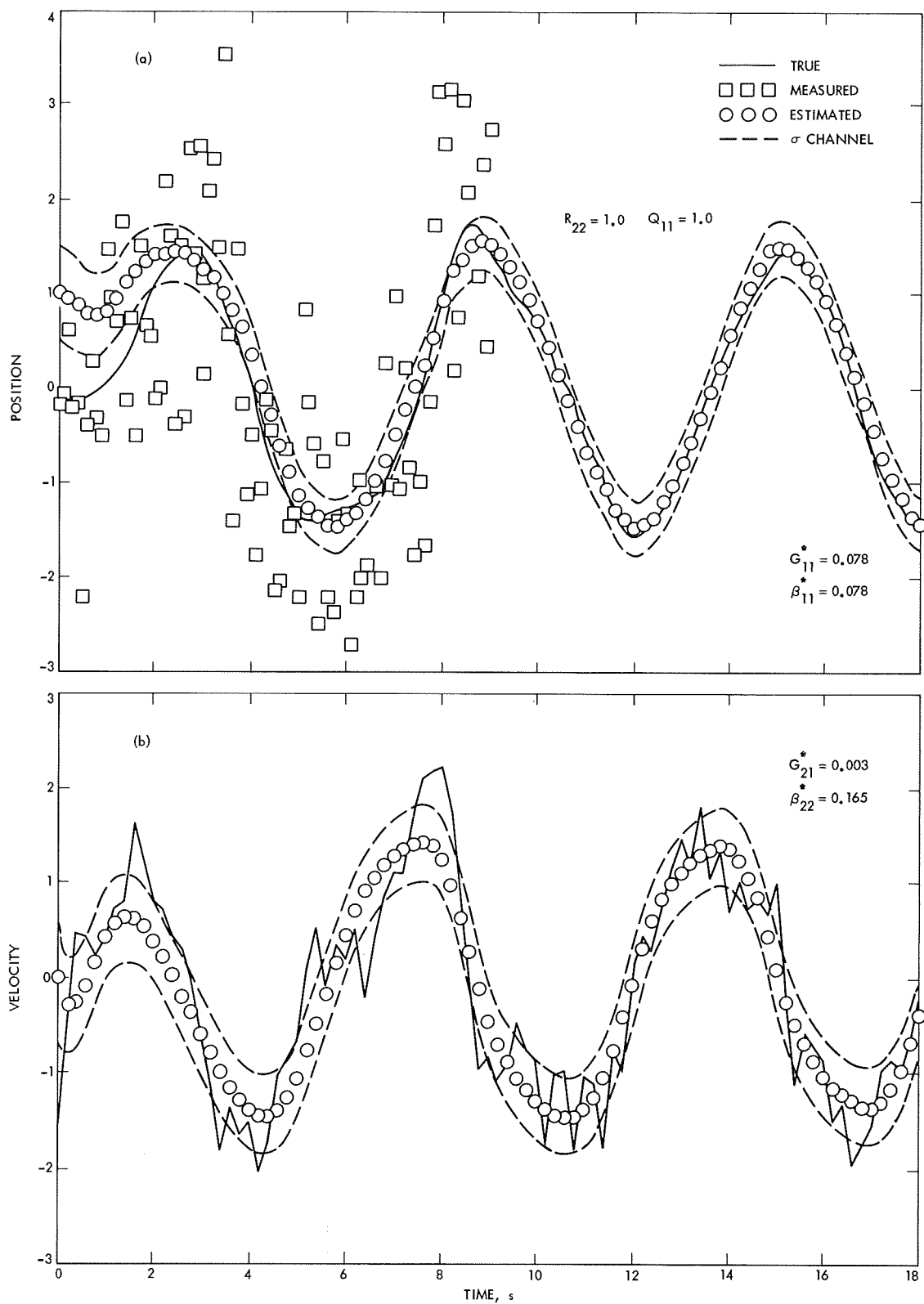


Fig. 17. A Monte Carlo sample case of a MSSVCG filter
for a nonlinear spring ($R_{22} = 1.0$, $Q_{11} = 1.0$)

E. Inertial Sensor Sterilization, P. J. Hand

As part of a continuing JPL series of evaluations of inertial sensors that might be capable of thermal sterilization, an Autonetics Type A-100 electromagnetic accelerometer (EMA) was subjected to an extensive stability test period, followed by six cycles of thermal sterilization at 135°C. A period of 164 days elapsed during these evaluations and 139 sets of data points were taken. In addition, the accelerometer was subjected to the JPL shock requirement of 200 g, 0.7 ms peak terminal saw tooth for a total of 15 shocks without significant damage.

The Autonetics accelerometer is a digitally rebalanced, linear, pulse-on-demand system of extremely small size. The acceleration sensor itself is a torque-balance, flexure suspended, pendulous, single-axis device. Pendulum position is obtained from a capacitive pickoff. All damping is obtained from the digital rebalance loop. The torquer is a fairly conventional moving coil design. However, the suspension system is quite unique in that the flexure, moving pickoff plate, and torque coil supports are all made from a single circular disc of fused quartz. The quartz is chemically etched to provide a cutout, which leaves a remaining area between the inner circular proof mass and the external support ring. This area is then additionally etched to reduce the thickness to 15–20 μm , providing a very compliant flexure. Electrodes are plated across this flexure to provide connections to the torque coil and pickoff. At the same time, the pickoff moving plate is plated directly on the quartz disc. This sensor section is housed in a miniature cylindrical container 1.125 in. in diameter by 0.65 in. high. The electronics required for the digital rebalance loop are housed in another cylindrical container 1.25 in. in diameter by 0.93 in. high. The electronics and sensor are combined along with a circular mounting flange and a connector, providing a package of 1.375 in. in diameter by 1.85 in. high. Total weight is 3.5 oz.

Additional electronics required to power and excite the digital rebalance loop are ± 24 Vdc, ± 12 Vdc, a 10.000 kHz clock signal and a 25.0 kHz pickoff and demodulator excitation source. All are external to the instrument and must be provided by the user. Clock and pickoff frequencies can be varied over a limited range, depending on the output scale factor required. Total input power to the instrument is 2.2 W.

For purposes of obtaining the daily stability characteristics of this instrument, it was mounted in a temperature-controlled fixture on an Ultradex dividing head. Output data was obtained in digital form and was recorded on

a Wang Laboratories Type 2003 bi-directional counter. A period of 100 s was used to accumulate the output count. Daily testing, which covered an initial period of 46 days, was largely confined to measuring the bias at zero-g input (B_0); and the positive and negative scale factors $K_1(+)$ and $K_1(-)$ at both room temperature and at 150°F. In addition, several test runs were made by measuring the above parameters at five temperatures between room temperature and 170°F. Two ± 1 -g linearity runs were also performed. The results of this pre-sterilization stability testing are shown in Figs. 18 and 19. After the first two weeks, a B_0 aging slope of $+1.0 \mu\text{g}/\text{day}$ is evident.

Thermal sterilization consisted of a 64-h exposure to a temperature of 135°C in a Tenney temperature chamber, with a gradual increase in temperature over a 4-h period prior to reaching 135°C. A similar time period was used on the cooling-down phase. B_0 and K_1 data were taken before and after each of the six sterilization cycles. The accelerometer was not operating during the

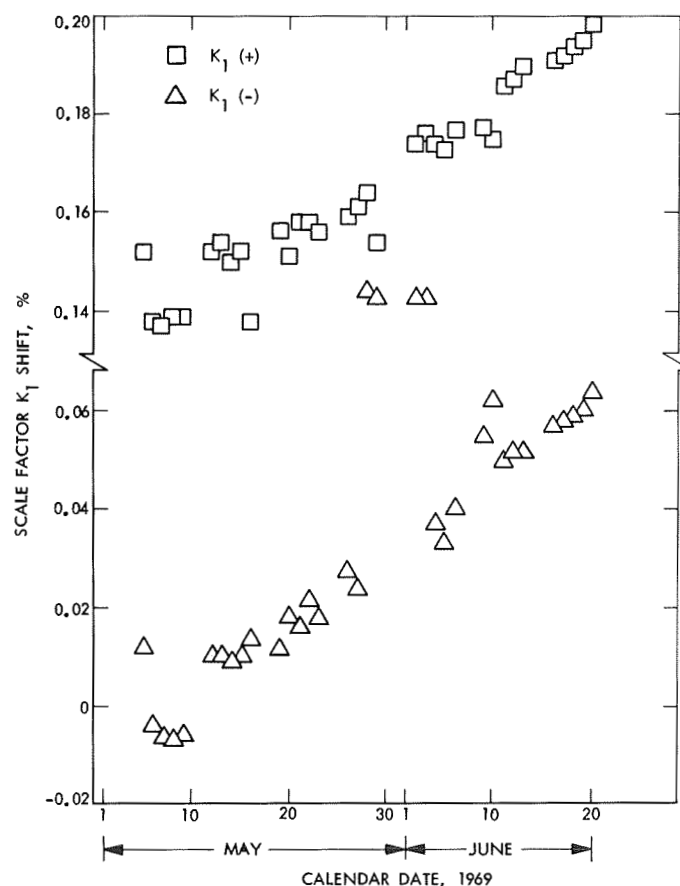


Fig. 18. Scale factor versus time—pre-sterilization (data taken at room temperature)

sterilization cycles. Shifts of the B_0 and $K_1(+)$ and $K_1(-)$ parameters are shown in Table 2. It will be noted that while the B_0 reading was fairly stable, the K_1 readings shifted by very large amounts with no indication of reaching a stable value.

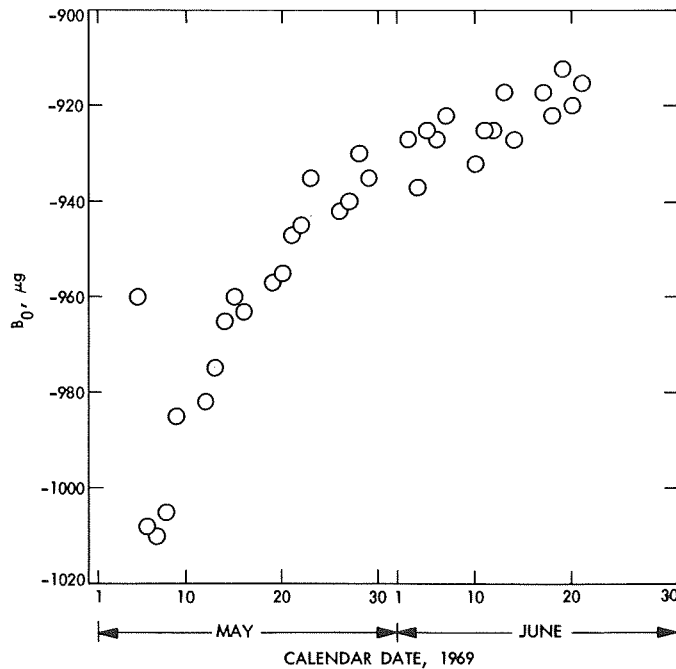


Fig. 19. B_0 versus time—pre-sterilization (data taken at room temperature)

Table 2. Parameter shifts from sterilization

Cycle	B_0 shift, μg	$K_1(+)$ shift, %	$K_1(-)$ shift, %
1	-254	+1.710	+1.761
2	+151	+0.277	+0.301
3	-153	+0.370	+0.277
4	+44	+0.111	+0.033
5	-83	+0.081	+0.030
6	+149	+0.178	+0.088
Mean shift per cycle	139 ^a	+0.455	+0.415

^aAbsolute value.
NOTE: Input axis was up during first cycle, down during second, up during third, etc., causing - and + variations on B_0 shift values.

Daily stability was again measured over a 47-day period after the sixth sterilization cycle. The results of this testing are shown in Fig. 20. Considerable loss of stability will be noted as well as the $K_1(+)$ and $K_1(-)$ stability values becoming noticeably poorer.

Exposure to 200-g, 0.7-ms shock is a typical type approval qualification requirement for JPL flight equipment. This instrument was shocked a total of 15 times (five shocks in each of three axes) without catastrophic failure. The B_0 shift as a result of the shocks was $-53 \mu g$ while the K_1 values changed $+0.103\%$ for $K_1(+)$ and $+0.128\%$ for $K_1(-)$. However $K_1(+)$ and $K_1(-)$ both returned to the original values within one day.

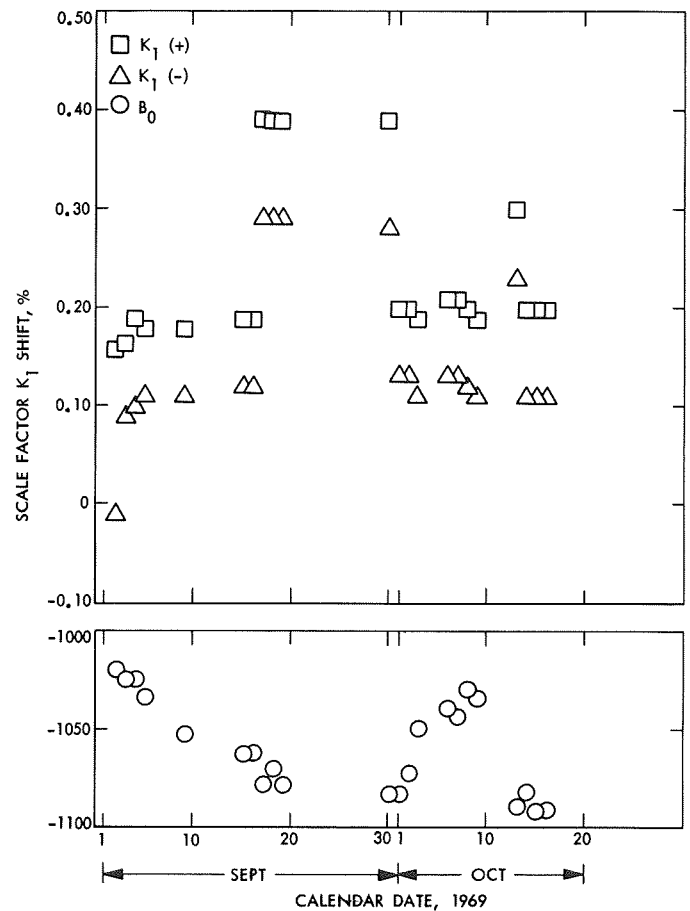


Fig. 20. Stability versus time—post-sterilization (data taken at room temperature)

XII. Guidance and Control Research

GUIDANCE AND CONTROL DIVISION

A. Cesium Breakdown Voltage in Thermionic Diodes, *K. Shimada*

1. Introduction

Volt-ampere curves of cesium-filled thermionic diodes operated at relatively low emitter and cesium reservoir temperatures clearly exhibit both unignited (low current) and ignited (high current) modes. The transition between the two modes occurs when cesium gas breaks down in the interelectrode gap. It is important to understand the physics of cesium breakdown due to volume ionization, a process which creates ions that reduce plasma losses in a thermionic converter.

Prior to cesium breakdown, the diode current increases with voltage in a manner that indicates an avalanche multiplication of electrons and ions. The formation of a visible glow on the surface of the collector during this process verifies that this mode is similar to the well-known anode-glow mode. At cesium breakdown the glow spreads throughout the interelectrode gap, and the diode current increases by several orders of magnitude.

Cesium breakdown voltages were measured in fixed-gap and variable-gap diodes as functions of emitter, collector, and cesium reservoir temperatures, for an

interelectrode distance D from 0.2 to 1.5 mm. Breakdown voltages varied from 5.0 to 1.7 V over the range of cesium pressure P from 0.0045 to 7 torr, but the voltage did not vary significantly with emitter and collector temperatures for any particular value of P . On the other hand, the interelectrode gap D affected the breakdown voltage considerably, so that the breakdown voltage was not the unique function of $P \times D$ that one would expect from a Paschen-type relationship. We suggest that the breakdown is, in fact, initiated in a region of the gap near the collector where a large portion of the breakdown voltage drop occurs, the remainder taking place in a dark plasma produced by electrons and ions from the emitter.

2. Experiments

Cesium breakdown voltages were measured in two guard-ringed converters, one of which had a fixed interelectrode gap of 10 mils (0.25 mm), the other a variable-gap between 0 and 60 mils (1.5 mm). Guard-ring potentials were electronically adjusted to follow the potentials of the collector so that the leakage currents between the guard-ring and the collector and spurious currents originating from parts other than the emitter could be eliminated during the measurements of volt-ampere characteristics.

The emitter temperatures were measured by an optical pyrometer aimed at a black body hole that was provided in the emitter, its temperature maintained within 5°K of the preselected values whenever the constancy of temperatures was important.

3. Fixed-Gap-Diode Results

This diode has basically the same dimensions as the solar energy thermionic (SET) converters. It has a planar electrode configuration with a collector (area 1.53 cm²) surrounded by a narrow guard ring. The guard ring is so constructed that its temperature stays nearly equal to that of the collector without need for the separate guard-ring heater that is required in conventional research converters. The materials for this diode¹ are tantalum for the emitter and niobium for both the collector and the guard-ring. The interelectrode gap is fixed at 0.25 mm at operating temperatures.

For measurements of the cesium breakdown voltages, the diode was operated at cesium pressures between 0.0045 and 7.0 torr, and true emitter temperatures between 1290 and 1708°K. The breakdown voltage V'_B is shown as a function of the cesium pressure in Fig. 1. V'_B is the sum of the measured breakdown voltage and the contact potential between the emitter and the collector and, hence, is the true potential difference in the interelectrode gap at breakdown. To arrive at a contact

potential, which is the difference between emitter and collector work functions, the emitter work function was calculated from Richardson's equation by using the true emitter temperature and the current density just prior to cesium breakdown, and the collector work function was determined independently for a similar range of temperatures. Thus the voltage V'_B is the maximum potential difference in the gap since the calculated emitter work functions were, in most cases, higher than theoretically expected because of the effects of electron space-charge barriers.

The results shown in Fig. 1 indicated that the voltage V'_B decreased from nearly 5 V at the low cesium pressure of 0.0045 torr to approximately 1.7 V at 1 torr; however, it increased at higher cesium pressures. Such behavior is similar to that described by Paschen's curve for breakdown in inert gases. V'_B was not affected by the emitter temperature T_E at pressures below 1 torr, but was affected considerably above 1 torr and decreased rapidly as the emitter temperature was increased. For temperatures higher than 1800°K, the breakdown voltages were not easily measured because the transition from the unignited to the ignited modes was smooth.

The effect of collector temperature was also examined for collector temperatures between 496 and 733°K at a fixed emitter temperature of 1373°K. Measured breakdown voltages varied as much as 0.3 V, but the true breakdown voltage V'_B , which included the effect of contact potentials, varied half as much (0.15 V). This value was within the experimental error, and hence the effect of collector temperature on the true breakdown voltage V'_B was considered negligible. It should be pointed out that electrical oscillations occurred under certain temperature combinations, the data for which have been excluded, since the oscillations caused anomalous results in the breakdown voltages.

4. Variable-Gap-Diode Results

This diode has a planar-electrode geometry with a tungsten emitter and a molybdenum collector. It was provided with two sapphire windows for: (1) visual observation of the interelectrode gap, (2) measurement of the interelectrode gap by cathetometer, and (3) measurement of emitter temperatures. The emitter and the electron-bombardment heater were constructed as one integral unit that was evacuated and separated from the main converter section containing the cesium. This construction allowed the diode to be operated in a normal atmosphere in contrast to others which require vacuum environments.

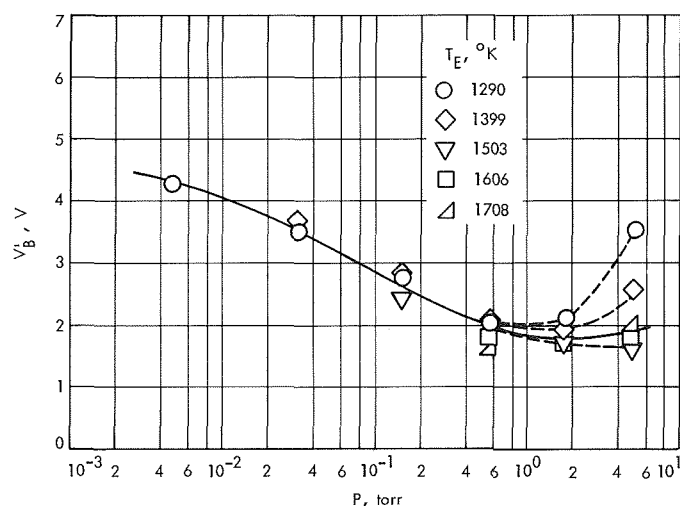


Fig. 1. Breakdown voltage in fixed-gap diode ($D = 0.254$ mm)

The breakdown voltage was measured as a function of interelectrode gap at fixed emitter temperature of 1273°K. During each run, the cesium reservoir temperature was kept constant within 1.0°K of a preset value which was varied between 453 and 553°K at 20°K intervals. The results are shown in Fig. 2. Note that the breakdown voltage V'_B was nearly constant for a low cesium temperature of 453°K, for which the cesium pressure was 0.03 torr; however, V'_B increased appreciably as the gap increased at the higher cesium temperature of 553°K, for which $P \approx 1.0$ torr. It should be pointed out that for $T_{Cs} < 510^\circ\text{K}$, the electron-neutral mean-free-path λ_e was such that $\lambda_e > D$ and that for $T_{Cs} > 510^\circ\text{K}$, $\lambda_e < D$.

To compare the relationship between V'_B and the pressure-distance product $P \times D$ with Paschen-type relationships, the results were replotted as shown in Fig. 3, in which it is clearly seen that V'_B is not a unique function of $P \times D$. Note here particularly that the curve shifts to larger $P \times D$ as well as to larger V'_B as D increases. The previous results obtained with a fixed-gap diode ($D = 0.25$ mm) are consistent with the present ones except that the previous curve falls on that for a gap D of approximately 0.37 mm, larger than expected. Possible reasons for discrepancy are that: (1) the electrode spacing of the fixed-gap device may have been larger at the operating temperature than the nominal value; (2) the emitter and the collector may not have been exactly parallel so that breakdown in the variable-gap diode could have occurred at lower voltages; and (3) the variable-gap diode may have had some of the so-called "pimples," which tend to form on electrodes when hot

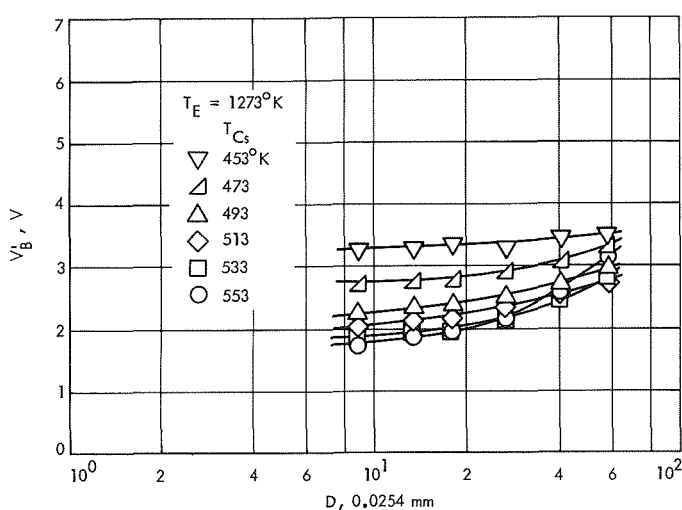


Fig. 2. Breakdown voltage in variable-gap diode ($D = 0.02$ – 1.5 mm)

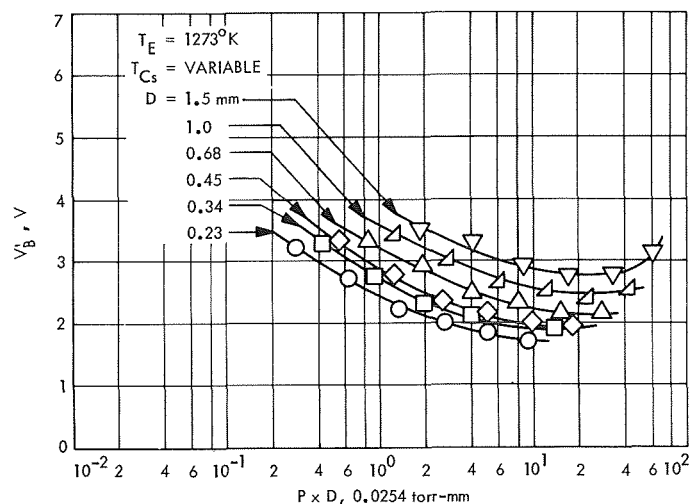


Fig. 3. Breakdown voltage versus $P \times D$

electrodes are brought into contact during the process of establishing zero-gap conditions, so that breakdown would occur prematurely.

Results such as those shown in Fig. 3 for which the breakdown voltage depends on D at fixed $P \times D$ may be explained phenomenologically by considering that: (1) the potential distribution is not linear in a cesium thermionic diode, and (2) breakdown is initiated in a region near the collector that is relatively narrow compared to the total gap D . The latter fact was substantiated by observing an anode glow in a relatively thin region adjacent to the collector just prior to cesium breakdown. Excess voltages ΔV required for breakdown at larger values of D were calculated from differences between values of V'_B at large D and those at $D = 0.23$ mm, which was taken as a reference gap for convenience. The results plotted in Fig. 4 show that ΔV increases with gap D and with the cesium pressure almost linearly.

5. Conclusion

The breakdown voltage of cesium in thermionic energy converters varies between 5.0 and 1.7 V as a function of cesium pressure between 0.0045 and 7 torr. The voltage does not vary significantly with the emitter temperature between 1300 and 1700°K, so long as such anomalies as electrical oscillations in the diodes do not occur. The collector temperature affected the observed breakdown voltage, but the true breakdown voltage V'_B was unaffected. The relationship between the voltage V'_B and the pressure-distance product $P \times D$ was analogous to that in the Paschen curves, except that V'_B was higher for larger D at a given $P \times D$. A plausible explanation is

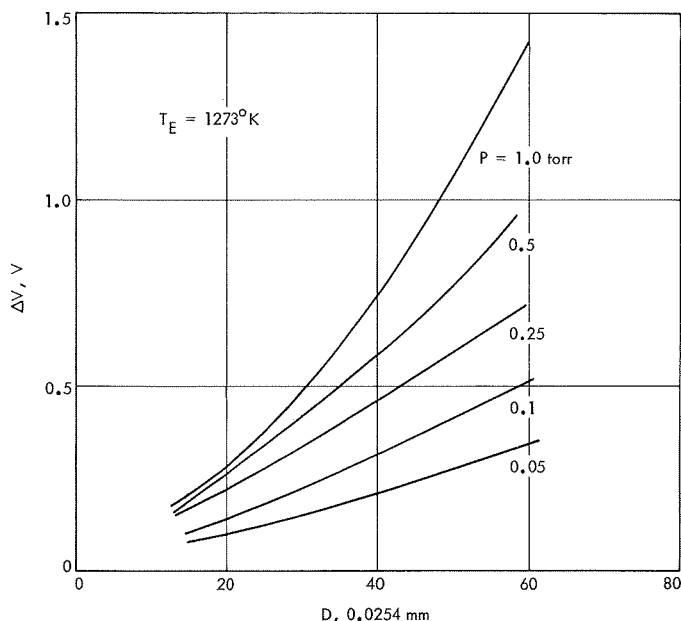


Fig. 4. Voltage drop in dark plasma

that breakdown is initiated in a narrow gap adjacent to the collector where the characteristic breakdown voltage drops occur, and that the remainder of the observed breakdown voltage occurs across the remainder of the interelectrode gap. The latter drop appears across the dark plasma produced by electrons and ions from the emitter. A theory that describes the dependence of diode current on voltage in regions including prebreakdown and breakdown regimes is now being developed. Full understanding of these regimes is important for reducing plasma losses in thermionic converters, especially those that would operate at lower temperatures.

B. Narrow Curie-Point Switching Transfer Characteristics in MnBi Films, G. Lewicki and J. Guisinger

1. Introduction

Thin ferromagnetic films of manganese bismuth each having an axis of magnetization perpendicular to the plane of the film are currently under investigation as media for magneto-optic information storage, with laser Curie-point switching as a means for writing in information and the Faraday effect for reading it out (SPS 37-48, Vol. III, p. 72). Previous investigations of Curie-point switching of MnBi films have shown the range over which the magnetic field controls switching is discouragingly large, on the order of 300 to 500 Oe (SPS 37-56, Vol. III, p. 145). A recent development in terms of practical application of Curie-point switching is the finding

that MnBi films can be prepared such that the range of field controlling Curie-point switching is on the order of 30 Oe and sometimes as small as 18 Oe. This information was gained by careful observation of the effect of various processing parameters involved in the preparation of the films on the Curie-point-switching characteristics of the films.

Curie-point switching is a technique used to control the direction of magnetization within preselected areas of the film, this direction representing information stored on the film. In laser Curie-point switching, a laser beam is focused onto the film to momentarily raise the temperature of an area above the Curie point, where the material becomes nonmagnetic. The area is then allowed to cool in the presence of an applied magnetic field which determines the final value and direction of the average magnetization within the area. This applied field does not affect the magnetization within areas of the film that have not gone through the heating-cooling cycle.

It is true that an area surrounded by a magnetically saturated region will, upon cooling from above its Curie temperature, acquire a magnetization opposite in direction to that of the surrounding film when there is no field applied during switching. But, to change the state of that area once again, that is, to switch its magnetization back to a direction parallel to that of the surrounding film, Curie-point switching has to be carried out in an applied magnetic field.

Inasmuch as the rewriting capability just described is an essential part of any magneto-optic information-storage scheme, it seems probable that a prime requirement for the practical application of MnBi films is that the range of applied magnetic field controlling Curie-point switching be small, since the generation of large magnetic fields varying at high frequencies over large areas is, at best, difficult.

Figure 5 contains a typical characterization of a newer film with a thickness of approximately 500 Å, a transmission of 16%, and a Faraday rotation of ± 4 deg at the helium-neon laser wavelength of 6328 Å. The M-H loop represents a plot of the average magnetization perpendicular to the plane of the films as a function of magnetic field applied perpendicular to the film without laser heating. It shows that, when magnetically saturated, the film remains saturated until a critical field in a direction opposite to that of the magnetization is reached. The critical field, usually called the coercive force, is 625 Oe for the film characterized in Fig. 5.

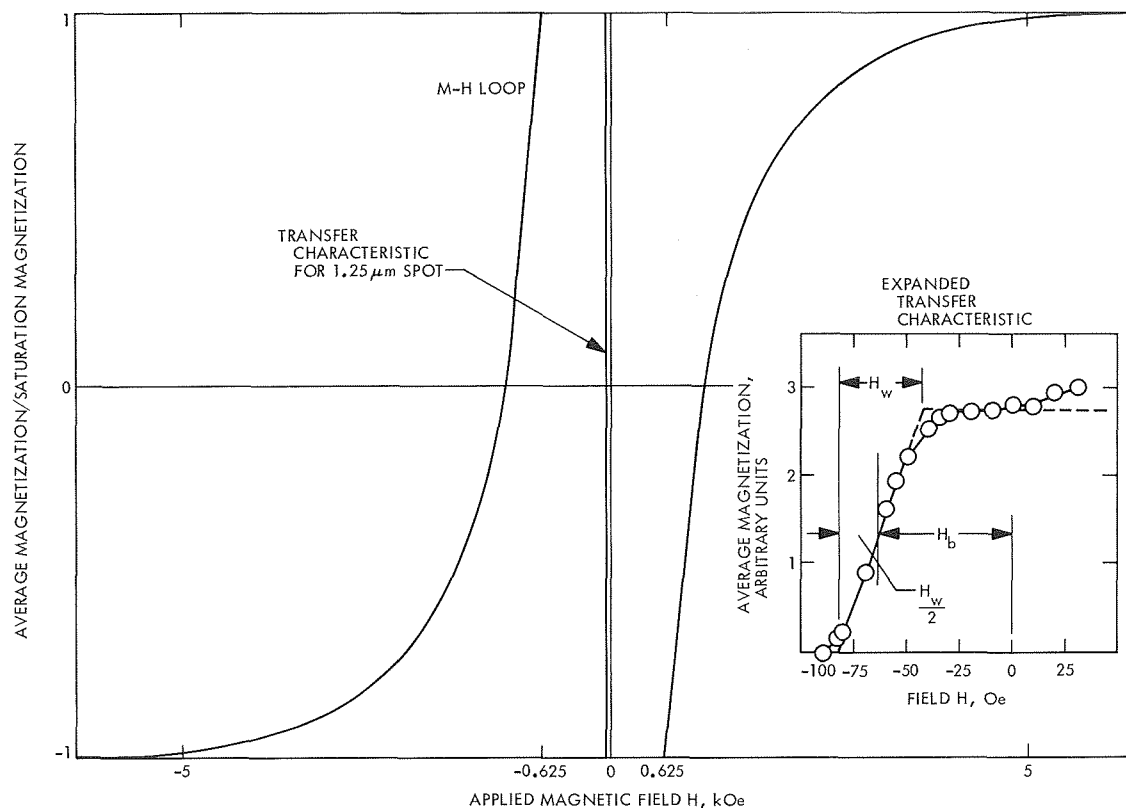


Fig. 5. Curie-point switching transfer characteristic of an MnBi film

The almost vertical line shown in Fig. 5 is the Curie-point-switching transfer characteristic for a spot with a diameter of $1.25 \mu\text{m}$. The transfer characteristic is a plot of the final value of the average magnetization within the spot following Curie-point switching as a function of magnetic field applied during switching, under the condition that the surrounding film is magnetically saturated with an average magnetization-to-saturation-magnetization ratio of -1 . All magnetizations in Fig. 5 were measured optically by sensing the rotation of plane-polarized light upon passing through the film.

An expanded plot of the transfer characteristic is shown in the insert in Fig. 5. The characteristic shows that if a $1.25\text{-}\mu\text{m}$ -diameter spot surrounded by a magnetically saturated region with magnetization $-M_s$ is Curie-point switched in an applied field equal to or smaller than $-H_b - H_w/2$, the final average magnetization will be approximately $-M_s$; if the applied field is equal to or greater than $-H_b + H_w/2$, the final average magnetization within the spot will be near $+M_s$. H_w is the range of field controlling Curie-point switching, and is equal to approximately 30 Oe in the characteristic shown in Fig. 5. Another requirement for the practical application of MnBi films is that areas of the film be able to take

sufficiently large excursions in temperature above the Curie point without permanent damage, to allow the size of Curie-point-switched spots to be relatively independent of the power of the beam. Immediately following

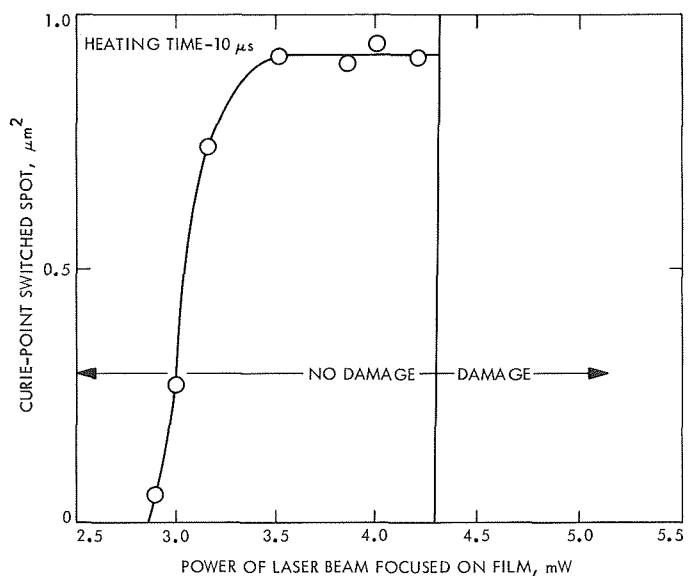


Fig. 6. Area of switched spot as a function of beam power used for heating (heating time $10 \mu\text{s}$)

laser heating, an area will have a temperature-position profile with an approximately Gaussian shape. The temperature will be highest at the center of the area, there varying slowly with position, and dropping off rapidly at the edge.

If the peak temperature is only slightly above the Curie temperature, the size of the area above the Curie temperature will be strongly dependent on the actual value of peak temperature; if the peak temperature is significantly above the Curie temperature, this size will be relatively independent of peak temperature.

A characterization of the newer MnBi films relevant to the considerations given above is shown on Fig. 6, where the area of a Curie-point-switched spot is plotted

as a function of the power of the laser radiation focused onto the film by a 0.32 numerical-aperture lens for a period of 10 μ s to heat the spot. It is seen that sufficiently large beam powers can be used to allow the size of Curie-point-switched areas to be independent of power but not damage the film.

The following areas will be researched: (1) a study of fatigue problems (if any) associated with Curie-point switching in the newer films, and (2) simple recording experiments in which tracks will be Curie-point switched by a focused laser beam moving across the film. During the scan, the magnetic field controlling Curie-point switching will be varied to effect a variation in the average magnetization and thus the magneto-optic density along the track.

XIII. Materials

ENGINEERING MECHANICS DIVISION

A. Tape-Type Temperature Control Coatings Application, J. C. Lewis

1. Introduction

The major proportion of temperature control on an unmanned spacecraft is accomplished through the use of various temperature control coatings. These coatings control the flow of heat at the surface in proportion to their absorptance, reflectance, and emittance.

In addition to the theoretical optical properties, the temperature control performance of a coating is affected by the application process. The quality of contact with the thermal mass and the uniformity of optical properties as applied to a complex component contribute to "effective" optical properties that often differ considerably from theoretical values.

Most temperature control coatings are "passive," i.e., the optical properties cannot be varied remotely in a controlled, predictable manner. Therefore, the ideal design condition is one where the optical properties can be determined prior to application, controlled during application, and retained after application.

This program was conducted to evaluate one temperature control material application method that showed potential for meeting these criteria. This method consists

of applying the temperature control coating in the form of pressure-sensitive tapes.

2. Review of State of the Art

Temperature control materials in tape form have been flown on a few satellites and spacecraft. In all the known cases, the adhesive was a pressure-sensitive silicone. Table 1 lists some of the uses of temperature control tapes known at this time.

Both McDonnell Douglas Corporation and General Electric Company currently have various tapes in vacuum and simulated space environment storage tests. Results of these tests have not been published yet.

3. Test Procedure

a. Mission-environment simulation. A thermal-vacuum test cycle was selected to simulate the severe but potentially real conditions such as sun occultation at Mercury. The range of selected temperatures was $400 \pm 25^\circ\text{F}$ to below -200°F . These temperatures represent the high temperature at equilibrium while exposed to the sun and cryogenic quenching during sun occultation.

Table 2 gives the details of the test sequence. No attempt was made to evaluate the effect of ultraviolet and particulate radiation on the tapes since the primary

Table 1. Use of pressure-sensitive temperature control tapes on spacecraft^a

Manufacturer	System	Type of exposure	Type of tape	Environment	Results
McDonnell Douglas Corp.	Gemini adaptor	Flight	Gold on Kapton	Space vacuum only	No apparent problems ^b
	Gemini adaptor	Flight	Aluminum foil	Total space environment	No apparent problems ^b
Lewis Research Center, NASA	Spacecraft electric rocket test	Type-approval test	Aluminum foil	Thermal-vacuum chamber	Blistering at 150° F on 24-in.-wide tape; no blistering on 4-in.-wide tape
General Electric Co.	Rocket motor tests in earth orbit	Flight	Aluminum foil	Total space environment	No apparent problems ^b
	Electronic packaging	Flight	Various tapes	Space vacuum only	No apparent problems ^b
North American Rockwell Corp.	Apollo 2 TV-1	Thermal-vacuum test	Aluminized Kapton	Thermal-vacuum chamber	Small gas bubbles; slight crinkling
	Apollo 5	Flight	Aluminized Kapton	Total space environment	No problems attributable to tape form
	Apollo 7	Flight	↓	↓	↓
	Apollo 8	Flight			
	Apollo 9	Flight			
	Apollo 10	Flight			

^aPersonal communications.
^bSpecific effects were not determined since recovery was not made.

objective was to evaluate the application interface. Extensive radiation stability studies are being conducted in various laboratories and were considered beyond the scope of this investigation. In addition, it is believed that such radiation primarily affects the temperature control coating material rather than the film-adhesive-substrate interfaces.

b. Equipment. An internally heated thin aluminum platen, 12 in. square, was fabricated by sandwiching a thin heating element between two 0.063-in.-thick aluminum plates (Fig. 1). This type of platen was chosen to simulate surfaces typical of spacecraft hardware. The temperature of the platen was internally monitored by 10 copper-constantan thermocouples placed in machined grooves on the inside surface of each plate. Five thermocouples were placed on each side of the platen. The thermocouples were located at the center of each plate and at the centers of each of the four quadrants of each plate.

The heating element was a 12-in.-square patch heating element which developed 10 W/in.² at 115 V and 12.5 A. Maximum operating temperature was 450°F. The heating

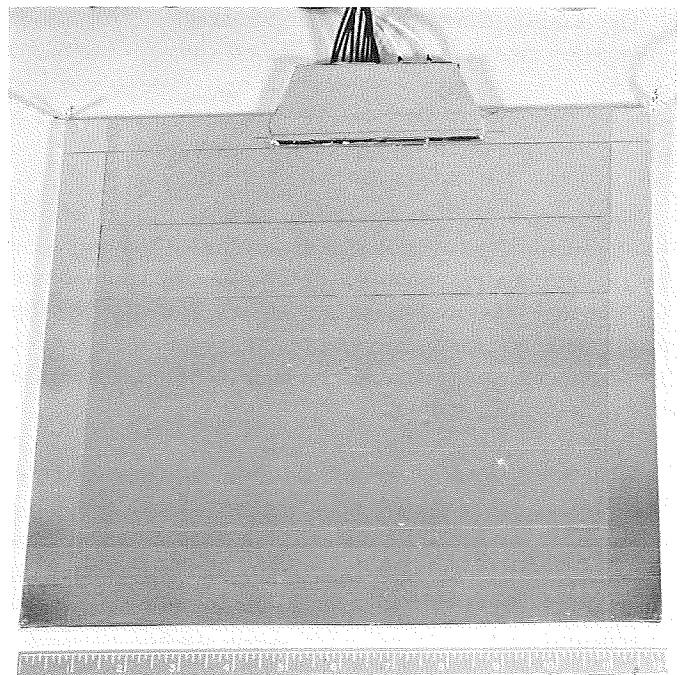


Fig. 1. Internally heated aluminum platen with aluminum-Teflon laminated tape prior to test

Table 2. Typical test sequence for simulated sun occultation at Mercury^a

Step	Operation	Data taken
1	Place platen in vacuum chamber with a shroud capable of liquid nitrogen (LN ₂) cooling and decrease pressure to less than 10 ⁻⁶ torr.	
2	Simultaneously cool LN ₂ shroud to -300 ⁺⁰ ₋₂₀ °F while applying sufficient power to maintain platen at ambient temperature. Allow platen to stabilize at ambient temperature with less than 2°F change in 1 h.	Temperature, voltage, current
3	Increase power to platen to achieve a temperature rise of approximately 100°F/h, stabilizing at 400 ±25°F.	Temperature, voltage, current
4	Cut off power to platen, allow platen to quench below -200°F.	Temperature
5	Increase power to platen to achieve a temperature rise of approximately 100°F/h until the power applied in Step 3 is applied. Allow platen temperature to stabilize at this power setting.	Temperature, voltage, current
6	Decrease power to the power setting applied in Step 2. Allow platen to stabilize at ambient temperature.	Temperature, voltage, current
7	Repeat Steps 4 and 5 at least five times.	Temperature, voltage, current
8	Repeat Step 6.	Temperature, voltage, current
9	Allow shroud to warm to ambient temperature, pressurize chamber to atmospheric pressure, remove platen from chamber.	
^a Chamber pressure maintained at less than 10 ⁻⁶ torr and LN ₂ shroud maintained at -300 ⁺⁰ ₋₂₀ °F for all testing.		

these methods was verified by the reproducibility of the applied energy-temperature data as seen in Tables 3, 4, and 5.

Table 3. Thermal analysis of test on aluminum foil-Teflon laminated tape^a

1969 date	Time	Circuit voltage, V ^b	Platen current, A	Platen resistance, Ω	Platen power, W	Platen temperature, °F ^c	Effective emittance ^d
3/31	09:40	22.54	2.70	8.25	60.1	92	0.62
	11:20	22.54	2.70	8.25	60.1	92	0.62
	14:35	51.04	6.10	8.27	307.6	376	0.60
	15:30	51.04	6.10	8.27	307.6	376	0.60
	15:31	0	0		0	376	
4/1	08:00	0	0		0	-255	
	13:05	51.00	6.10	8.27	307.6	376	0.60
	14:00	50.99	6.09	8.27	306.7	377	0.60
	15:55	22.49	2.69	8.26	59.8	96	0.60
	16:55	22.50	2.70	8.23	60.0	94	0.61
4/10	09:50	22.50	2.69	8.26	59.8	90	0.62
	10:30	22.50	2.69	8.26	59.8	92	0.62
	13:35	51.03	6.10	8.26	307.3	375	0.60
	15:00	51.06	6.10	8.27	307.6	375	0.60
	16:01	0	0		0	375	
4/11	08:00	0	0		0	-257	
	13:00	51.02	6.10	8.26	307.3	375	0.60
	15:55	51.08	6.10	8.27	307.6	377	0.60
	16:00	0	0		0	377	
4/14	08:00	0	0		0	-272	
	13:35	51.08	6.11	8.26	308.1	375	0.60
	15:15	51.02	6.10	8.27	307.6	375	0.60
	15:31	0	0		0	375	
4/15	08:10	0	0		0	-257	
	13:50	51.00	6.10	8.27	307.6	374	0.61
	16:00	51.04	6.10	8.27	307.6	375	0.60
	16:01	0	0		0	375	
4/16	08:00	0	0		0	-258	
	14:15	51.07	6.10	8.27	307.6	375	0.60
	16:00	51.05	6.10	8.27	307.6	375	0.60
	16:01	0	0		0	375	
4/17	07:00	0	0		0	-255	
	13:25	51.06	6.10	8.27	307.6	376	0.60
	16:00	51.07	6.10	8.27	307.6	376	0.60
	16:01	0	0		0	376	
4/18	04:10	0	0		0	-235	
	09:05	51.05	6.10	8.27	307.6	378	0.60
	10:10	51.00	6.09	8.27	306.7	378	0.59
	12:20	22.54	2.70	8.25	60.1	95	0.61
	13:25	22.56	2.70	8.26	60.2	94	0.61
^a Platen was removed from chamber on 4/1/69 and reinstalled unmodified on 4/10/69.							
^b Circuit voltage includes the voltage drop across 0.1-Ω resistance within the power supply.							
^c Temperature listed is the average temperature recorded for the 10 thermocouples.							
^d Calculations of effective emittance are based on 300 in. ² of radiating area and should not be taken as absolute values.							

element was purchased from Electrofilm, Inc., of North Hollywood, California, who installed the thermocouples and bonded the heater to the two plates.

By heating internally and covering both surfaces and all edges with the temperature control tape under investigation, all the energy applied to the heater can be assumed to be emitted by the tape since special efforts were taken to minimize conductive heat losses through the power and thermocouple leads. These methods included the use of leads approximately 6 ft long, which were wrapped with aluminized Mylar to achieve minimum emittance of conducted heat losses. The success of

Table 4. Stability of platen at equilibrium





Date and time	Thermocouple reading, °F									
	1	2	3	4	5	6	7	8	9	10
22.66 V and 2.64 A										
3/20/69										
11:25	84	87	87	87	84	79	84	82	83	80
11:30	84	87	87	87	84	79	83	82	83	80
11:35	84	87	87	87	84	79	84	82	83	80
11:40	84	87	87	87	84	79	84	82	83	80
11:45	84	87	87	87	84	79	84	82	83	81
11:50	84	87	87	87	84	79	84	82	83	80
11:55	84	87	87	87	84	79	84	82	83	81
12:00	84	87	87	87	84	79	84	82	83	81
Zero voltage and current										
3/21/69										
01:00	-223	-224	-223	-223	-224	-224	-223	-223	-224	-223
02:00	-229	-230	-230	-230	-230	-230	-230	-230	-230	-230
03:00	-234	-234	-234	-234	-234	-234	-234	-234	-235	-234
04:00	-238	-239	-239	-239	-239	-239	-239	-239	-239	-239
05:00	-242	-242	-242	-243	-242	-242	-243	-242	-243	-242
06:00	-244	-245	-244	-245	-245	-245	-245	-244	-245	-244
07:00	-247	-248	-248	-248	-248	-248	-248	-248	-248	-248
08:00	-249	-250	-249	-250	-250	-250	-250	-250	-250	-250
51.04 V and 6.10 A										
3/31/69	 Inoperative 									
14:35		390	389	389	382	360	368	366	368	363
14:40		390	389	389	382	360	368	366	368	363
14:45		390	389	389	382	360	368	366	368	364
14:50		390	389	389	382	360	368	366	368	364
14:55		390	389	389	382	360	368	366	368	364
15:00		390	388	389	382	360	368	366	368	363
15:05		390	389	389	382	360	368	366	368	363
15:10		390	388	389	382	359	368	366	368	363
15:15		390	388	389	382	359	368	366	368	363
15:20		390	388	389	382	359	368	366	368	363
15:25		390	389	390	382	360	368	366	368	364
15:30		390	389	389	382	360	368	366	368	363
51.08 V and 6.11 A										
4/14/69	 Inoperative 									
13:35		390	389	389	383	358	367	365	368	361
13:40		390	389	389	383	358	367	365	367	361
13:45		390	389	389	383	358	367	365	368	361
13:50		390	389	389	383	358	367	365	368	361
13:55		390	388	389	383	359	367	365	368	361
14:00		390	389	389	384	359	367	365	368	361
14:05		390	388	389	383	358	367	365	368	361

Table 4. (contd)

Date and time	Thermocouple reading, °F									
	1	2	3	4	5	6	7	8	9	10
51.08 V and 6.11 A (contd)										
14:10	<div style="display: flex; align-items: center; justify-content: center;"> <div style="text-align: center; margin-right: 5px;">↑</div> <div style="writing-mode: vertical-rl; transform: rotate(180deg);">Inoperative</div> <div style="text-align: center; margin-left: 5px;">↓</div> </div>	390	389	389	383	358	367	365	368	361
14:15		390	388	389	383	358	366	365	368	361
14:20		390	389	389	383	358	367	365	368	361
14:25		390	388	389	383	358	367	364	367	360
14:30		390	389	389	383	358	367	365	368	360
14:35		390	389	389	383	358	367	365	367	359
14:40		390	389	389	383	358	367	365	368	360
14:45		390	389	389	383	358	367	365	368	359
14:50		390	389	389	383	358	367	365	368	360
14:55		390	389	389	383	358	367	365	368	359
15:00		390	388	389	383	358	367	365	368	359
15:05		390	389	389	383	358	367	365	368	359
15:10		389	388	388	383	358	367	365	368	359
15:15		390	388	389	383	358	367	365	368	359

Table 5. Thermal analysis of test on aluminized Kapton tape

1969 date	Time	Circuit voltage, V ^a	Platen current, A	Platen resistance, Ω	Platen power, W	Platen temperature, °F ^b	Effective emittance ^c	1969 date	Time	Circuit voltage, V ^a	Platen current, A	Platen resistance, Ω	Platen power, W	Platen temperature, °F ^b	Effective emittance ^c
5/26	17:00	15.80	1.89	8.26	29.5	74	0.35	6/3	07:30	0	0		0	—209	
5/27	08:25	15.80	1.89	8.26	29.5	74	0.35		13:40	36.83	4.41	8.25	160.5	388	0.30
	13:40	36.82	4.41	8.25	160.5	385	0.30		14:55	36.83	4.41	8.25	160.5	388	0.30
	14:25	36.82	4.41	8.25	160.5	386	0.30		14:56	0	0		0	388	
	14:30	0	0		0	386									
5/28	07:40	0	0		0	—218		6/4	07:30	0	0		0	—226	
	12:15	36.82	4.41	8.25	160.5	385	0.30		13:45	36.82	4.41	8.25	160.5	384	0.30
	13:00	36.82	4.41	8.25	160.5	386	0.30		15:10	36.82	4.41	8.25	160.5	384	0.30
	16:45	15.82	1.89	8.27	29.5	82	0.33		15:11	0	0		0	384	
	17:55	15.82	1.89	8.27	29.5	81	0.33	6/5	07:30	0	0		0	—227	
	18:00	0	0		0	81			13:50	36.82	4.41	8.25	160.5	386	0.30
5/29	07:30	0	0		0	—210			15:15	36.82	4.41	8.25	160.5	386	0.30
	13:35	36.82	4.41	8.25	160.5	384	0.30	6/6	15:16	0	0		0	386	
	15:05	36.82	4.41	8.25	160.5	385	0.30		07:30	0	0		0	—219	
	15:06	0	0		0	385			13:10	36.82	4.41	8.25	160.5	384	0.30
6/2	08:25	0	0		0	—249			14:00	36.82	4.41	8.25	160.5	385	0.30
	15:45	36.82	4.41	8.25	160.5	387	0.30		17:30	15.80	1.89	8.26	29.5	80	0.33
	16:55	36.84	4.41	8.25	160.5	387	0.30		18:05	15.81	1.89	8.26	29.5	80	0.33
	16:56	0	0		0	387									

^aCircuit voltage includes the voltage drop across 0.1-Ω resistance within the power supply.^bTemperature listed is the average temperature recorded for 4 thermocouples corrected to average temperature for 10 thermocouples from previous tests.^cCalculations of effective emittance are based on 300 in.² of radiating area and should not be taken as absolute values.

The platen suspended in the chamber with the shroud door removed is seen in Fig. 2.

c. Materials. Two tapes were tested during the program. These tapes were a commercially available aluminum foil-Teflon film laminate with a silicone adhesive on the aluminum foil side and the aluminized Kapton tape used on the *Apollo* vehicles (Table 1).

The laminated tape consisted of 0.002-in.-thick aluminum foil with 0.001-in.-thick Type C Teflon adhesively bonded to the foil. The aluminized Kapton tape was 0.0005-in.-thick Kapton with 1200-Å minimum thickness of vacuum coated aluminum covered with 10,000-Å minimum thickness of SiO_x . Both tapes had pressure-sensitive silicone adhesives.

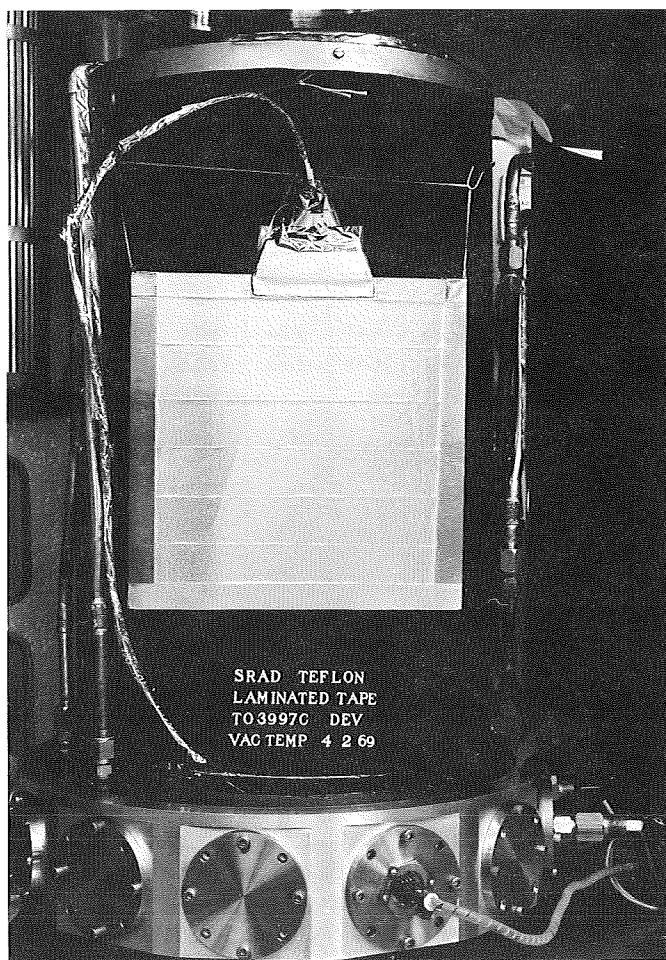


Fig. 2. Platen with aluminum-Teflon laminated tape suspended within LN_2 shroud in vacuum chamber

The actual temperature control materials on each of these tapes are known to degrade in ultraviolet radiation. However, as previously stated, we were interested in the behavior of the film-adhesive-substrate interfaces rather than the performance of the temperature control coating in the space radiation environments. The lack of availability of a commercial tape with a high-emittance coating that is stable in ultraviolet radiation prevented ideal simulation of what might be flown on a Mercury exploration mission.

In addition to being readily available commercially, these two tapes represented two extremes of film rigidity.

The aluminum-Teflon tape was purchased in 1-, 2-, and 25-in.-wide rolls. Only a 2-in.-wide roll of the aluminized Kapton tape was purchased. A 2-in.-wide roll of a commercially available tape consisting of gold over Kapton was also purchased. This tape was not tested because poor adhesion of the gold prevented overlapping of the tape during layup.

d. Degassing. The characteristic of silicones to outgas in vacuum required the platen and tapes to be baked at $400 \pm 25^\circ\text{F}$ for 24 h at less than 10^{-3} torr before use. The tapes were baked using the cylindrical drum shown in Fig. 3. Table 6 shows the efficacy of the bake out cycle.

The release paper had to be removed from the aluminum-Teflon tape prior to baking. Because Teflon was used as release film, the aluminized Kapton tape was baked with the Teflon release film in place. The thin Kapton film would have been much more difficult to handle and apply if the Teflon release film had had to be removed during baking.

e. Tape application. The tapes were applied to the platen by the author, who is not experienced in tape application. Therefore, the quality of the application does not reflect what is obtainable. Since the test was designed to show relative changes in thermal performance of the tape, it is believed that the wrinkles did not significantly affect the results of the tests.

The tapes were applied with a standard 4.5-lb roller recommended by the Pressure Sensitive Tape Council (Ref. 1). Additional pressure beyond the weight of the roller was used to smooth the tape and remove entrapped air bubbles. The air bubbles and wrinkles on the aluminized Kapton tape were punctured to release entrapped air.

Table 6. Outgassing characteristics of tapes

Tape	Tape Condition	Total weight loss, g/cm ²	Volatile condensible material ^a	
			Weight, g/cm ²	Description
Aluminum-Teflon laminate	As received	6.8×10^{-5}	1.2×10^{-5}	Not described
		4.2×10^{-5}	1.9×10^{-5}	Not described
	Baked at $400 \pm 25^\circ\text{F}$ for 24 h at less than 10^{-3} torr (backing paper removed)	4.6×10^{-5}	0.2×10^{-5}	Light cloudy deposit
		4.1×10^{-5}	0.2×10^{-5}	Light cloudy deposit
Aluminized Kapton	As received	5.2×10^{-5}	1.6×10^{-5}	Heavy oily deposit
		4.5×10^{-5}	1.1×10^{-5}	Heavy oily deposit
	Baked at $400 \pm 25^\circ\text{F}$ for 24 h at less than 10^{-3} torr (Teflon backing not removed)	1.3×10^{-5}	0.6×10^{-5}	Heavy oily deposit
		3.4×10^{-5}	0.8×10^{-5}	Heavy oily deposit

^aVolatile condensible material is defined as that material that condenses on a surface maintained at 25°C.



Fig. 3. Aluminized Kapton tape on drum after baking

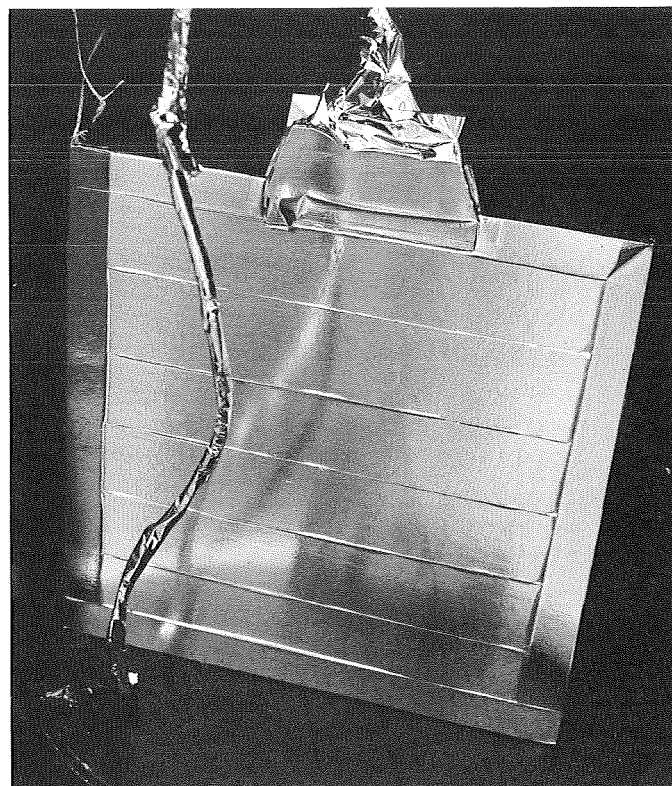
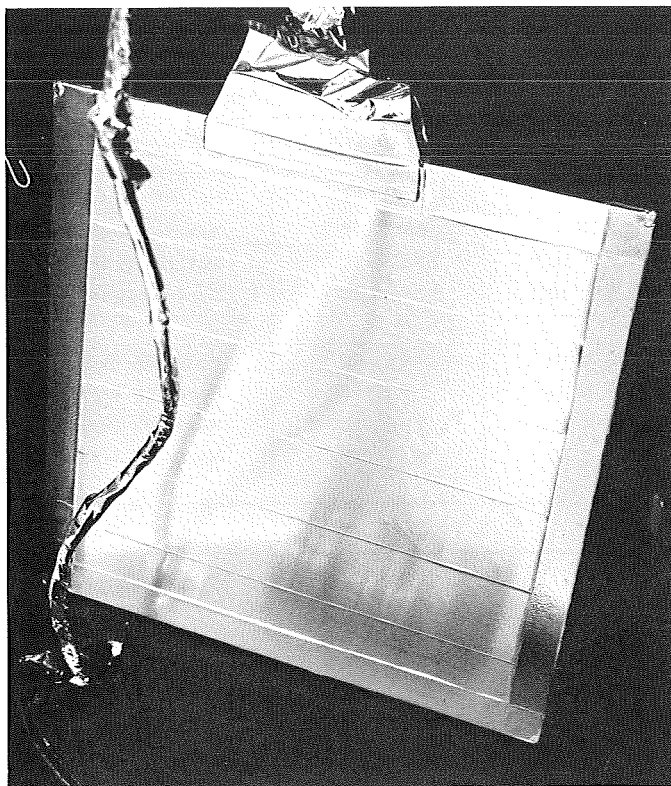


Fig. 4. Opposite sides of aluminum-Teflon laminated tape after test

The aluminum-Teflon tape edges were butted rather than overlapped because of poor adhesion of the pressure-sensitive silicone adhesive to the Teflon on the outer surface. The aluminized Kapton tape was overlapped.

4. Test Results

a. Aluminum-Teflon tape. Results of the test with the aluminum-Teflon tape are summarized in Table 3. The condition of the tape before test is shown in Fig. 1, while the condition of the tape after the test is seen in Fig. 4. Peel tests of the tape made in place on the platen showed 0.25 lb/in. of width "as applied" and 1.15 lb/in. after the test.

An analysis of Table 3 shows little change in the effective emittance between ambient temperature and 375°F. The small change that did take place is believed to be primarily due to a change in emissivity with temperature. The reproducibility of the effective emittance over nine cycles at 375°F is noteworthy.

The reliability, stability, and reproducibility of the heating platen were astonishing as can be seen from the data. A true rms digital voltmeter was used to monitor

voltage and current. This voltmeter permitted the voltage and current settings to be reproduced to the second decimal place, thereby eliminating any corrections for changes in applied power. The stability of the platen when stabilized can be appreciated from the thermocouple readings from typical equilibrium periods seen in Table 4.

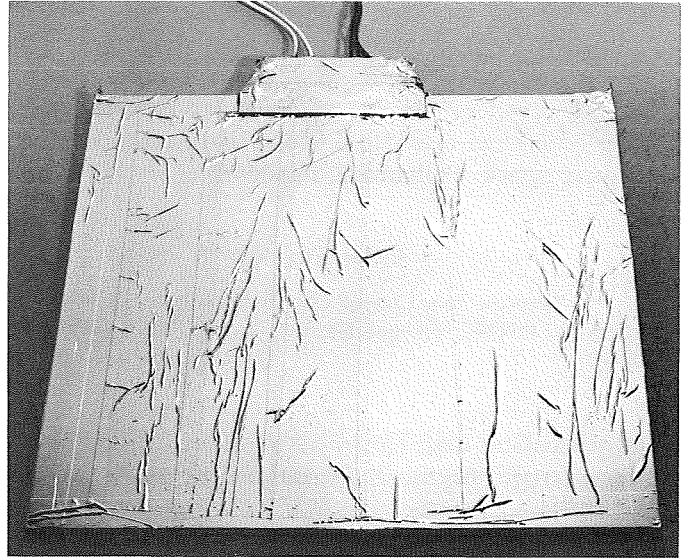
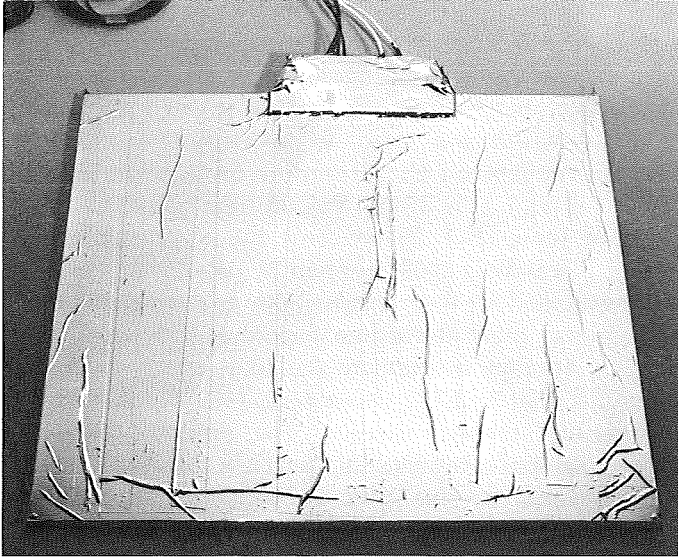
b. Aluminized Kapton tape. The thermal analysis of the test of the aluminized Kapton tape is shown in Table 5. It was observed after the test that the tape had a strong odor resembling charred wood. No visible effect was detected nor was the cause or the source of the odor determined. Comparison of photographs of the tape before and after the test showed little if any loss of contact between the tape and the platen (Fig. 5).

A peel test of one strip of tape showed a strength of approximately 0.3 lb/in. after the test. Additional peel tests could not be made because the tape tore in such a way that the full strip could not be gripped.

5. Conclusions

The tests conducted in this program indicate that pressure-sensitive tapes offer high potential as a base for

(a) OPPOSITE SIDES BEFORE TEST



(b) OPPOSITE SIDES AFTER TEST

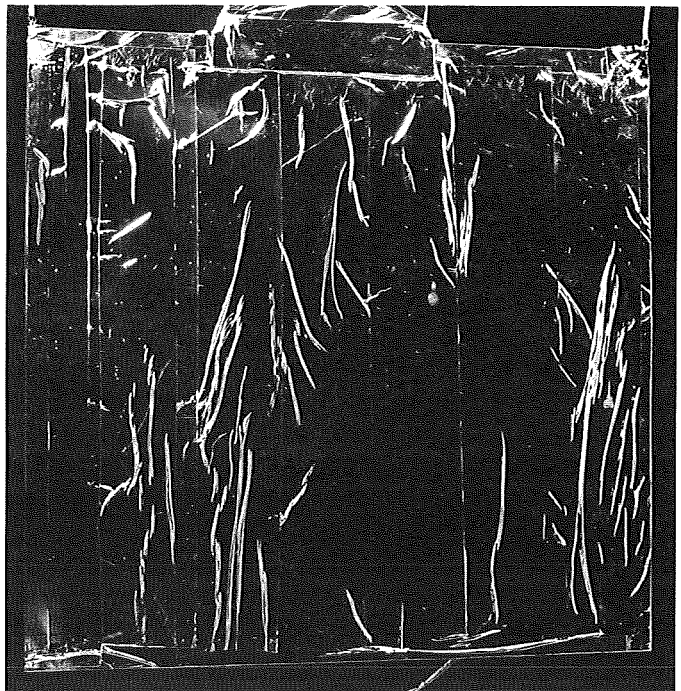
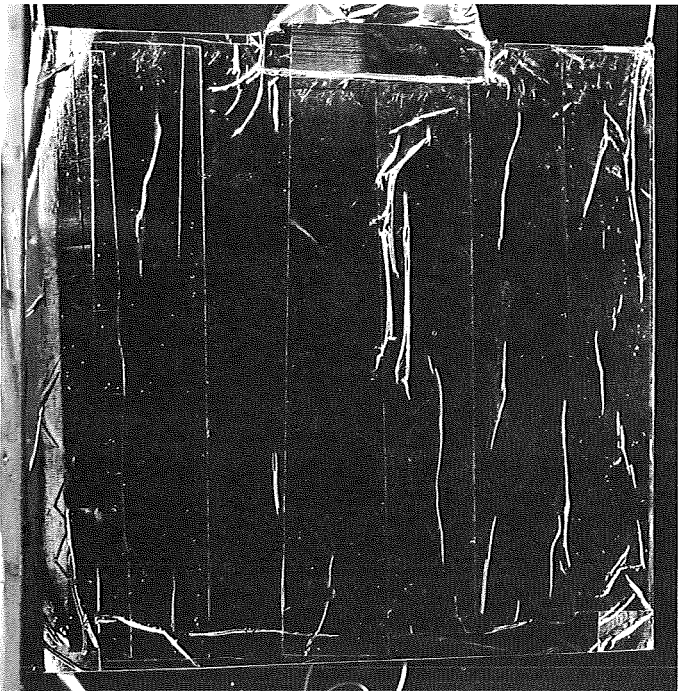


Fig. 5. Aluminized Kapton tape before and after test

temperature control coatings for spacecraft. The performance of such tapes is stable and predictable.

Where flexibility of application is not needed, polymeric films with thermosetting adhesives also could be used as bases for spacecraft temperature control coatings.

The intrinsic quality control and repairability obtainable with temperature control coatings in tape form justifies additional evaluation of this application method for future space exploration missions.

Reference

1. *Pressure Sensitive Tape Council 180° Peel Adhesion Test Specification*, No. PSTC-1. Pressure Sensitive Tape Council, Glenview, Illinois.

B. Sterilization and Thermal-Vacuum Effects on Spacecraft Polymeric Materials, W. Roper

1. Introduction

In future planetary entry/lander missions, it is anticipated that biological sterilization will be a necessary step in spacecraft preflight preparation. One of the major objectives of planetary exploration, the detection of extra-terrestrial life forms, would not be realized if viable earth-type organisms were carried aboard nonsterile spacecraft. A proposed preflight sterilization consists of decontamination by treatment with a gaseous mixture of Freon and ethylene oxide (ETO), and subsequent dry-heat sterilization in an inert nitrogen atmosphere.

All material used in spacecraft fabrication must, therefore, be compatible with sterilization processes. The sterilization requirement was thought to be particularly severe for many organic polymeric products (adhesives, films, tapes, etc.) used in the spacecraft. The amount of degradation will depend, in general, on the intrinsic chemical stability of the basic polymer and the various additives compounded into the final product.

In addition to the preflight sterilization, these same polymeric products must be capable of withstanding the anticipated space environment. This environment is defined as the thermal-vacuum conditions existing within a scientific satellite, or unmanned spacecraft, or within the unpressurized portions of a manned spacecraft. This environment is, therefore, one in which the polymeric materials have been essentially shielded from extremes of temperature and from electromagnetic and particle radiation. This environment has, therefore, generally

been considered to comprise a pressure of 10^{-5} torr or less and a temperature maximum of 135°C . The duration of exposure to this environment will obviously depend on the particular planetary mission; however, exposures of the order of 500 h are considered minimal.

Under NASA's Sterilizable Polymers Program, JPL has been evaluating, over the past few years, the effects of sterilization and thermal vacuum on selected polymeric products which are used in spacecraft fabrication. Several reports on these studies have been issued (Refs. 1-7 and SPS 37-40, Vol. IV, pp. 103-106). The latest of these (Ref. 7) concerns the work performed during Fiscal Year 1969, wherein 21 polymeric products were evaluated for their compatibility to sterilization plus a long-term thermal-vacuum exposure. The work was the first attempt to evaluate the total effect of all the exposures, that is, ETO decontamination, thermal sterilization, and thermal vacuum. The products included several material categories, such as films, adhesives, tapes, etc. Each product was given the following sequence of exposures:

- (1) Six decontamination cycles of a humidified ETO-Freon atmosphere at 50°C for 30 h each.
- (2) Six thermal-sterilizing cycles of a dry nitrogen atmosphere at 135°C for 96 h each.
- (3) One thermal-vacuum exposure for 500 h under conditions of 135°C temperature and 10^{-6} torr pressure.

2. Experimental Procedures

The polymeric products selected for evaluation in this recent program included the following material categories:

- (1) Adhesives—structural.
- (2) Coatings—conformal.
- (3) Coated fabrics.
- (4) Elastomers.
- (5) Encapsulants.
- (6) Films.
- (7) Hardware and structural materials.
- (8) Tapes.

The test sequence employed is outlined in Fig. 6. As indicated, materials testing was performed on control specimens and then on the exposed specimens in order

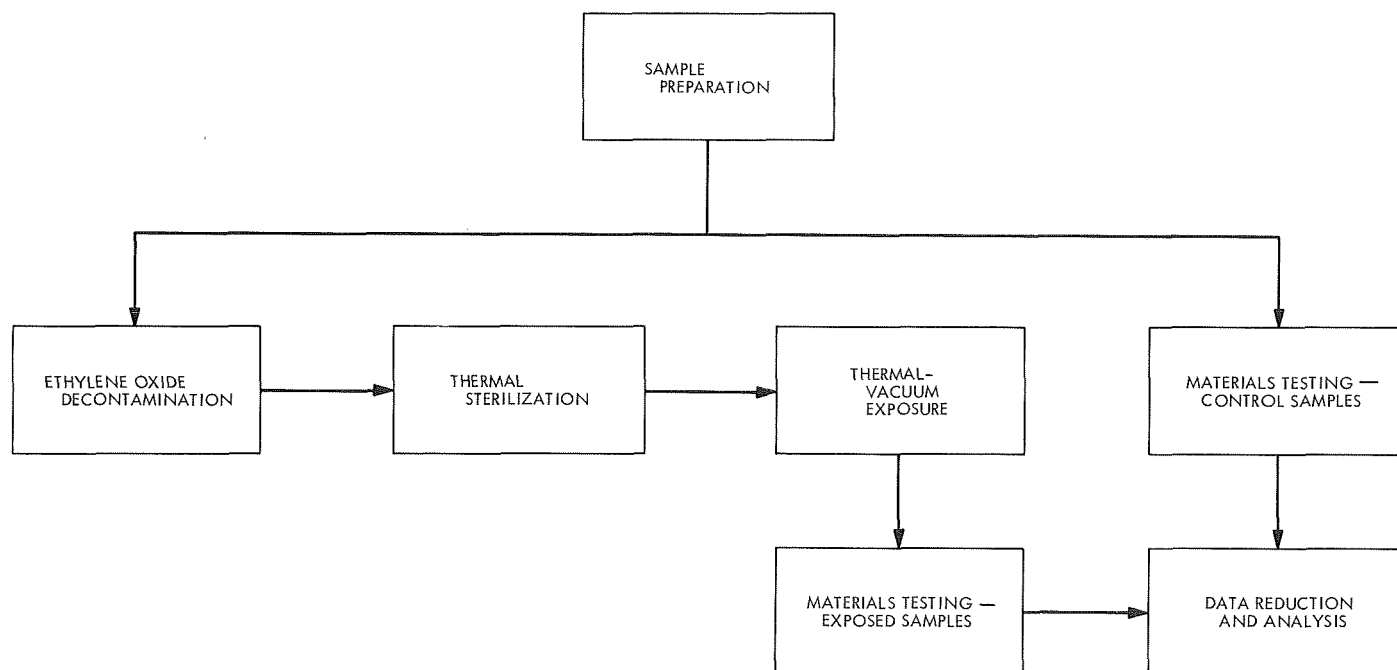


Fig. 6. Exposure and test sequence

to evaluate the effects of the exposures on the various products. The physical, mechanical, and electrical property tests that were employed for each class of materials are indicated in the materials test matrix shown in Table 7. The test specimens of each material were prepared in accordance with the sizes and shapes specified in the particular standard American Society for Testing Materials (ASTM) test method used. In most cases, a minimum of six specimens (three control and three exposed specimens) were prepared from each polymeric product.

The special exposure equipment used in the program included an ETO-Freon 12 decontamination chamber, a thermal sterilization chamber, and a thermal-vacuum exposure unit.

The equipment and procedures followed in the various exposures are given in detail in Ref. 7. The sterilization procedures followed were in accordance with JPL specifications. The equipment used is shown in Figs. 7, 8, and 9.

3. Results and Discussion

In order to establish whether a product was compatible with the various exposures, a rating scheme was necessary. In this program, the criteria and rating scheme used previously (described in Refs. 2 and 3) were employed.

In this scheme, the performance of each product is given a rating of compatible (C), marginal (M), or noncompatible (NC), according to the degree of retention of the key properties of the product after the various exposures. The key properties may be physical, mechanical, or electrical. The relative importance of key properties (e.g., mechanical versus electrical) depended on the particular material being evaluated.

The results of the study are summarized in Table 8. A detailed description of the test results may be found in the final report (Ref. 7). Of the 21 products evaluated, 55% were rated compatible, 30% marginal, and 15% noncompatible with the three exposures. It should be noted that it was not the intent of the work to separately evaluate the effects of the individual exposures. The total effect of all exposures was the sole objective.

4. Subsequent Studies

As a concluding task within the Sterilizable Polymers Program, a Materials Interaction Study is currently underway. In this program several sets of polymeric products have been selected for further study. These products are being placed in intimate contact and then given sequential exposure to the sterilization and thermal-vacuum environments. It is the purpose of the study to establish if any incompatibility between the products can be observed as a result of their contact and physical and

Table 7. Materials test matrix

Test	Adhesives	Coatings	Coated fabrics	Elastomers	Encapsulants	Films	Hardware and structural materials	Tapes	ASTM standard
Adhesive shear strength	X								D1002
Adhesion	X	X						X	D1000 D1876 D2197 D2294
Adhesive creep	X								—
Weight change	X	X	X	X	X	X	X	X	—
Volume resistivity		X			X	X	X		D257
Dielectric strength (short time)		X			X	X	X	X	D149 D1000
Breaking strength			X					X	D1000 D751
Dimensional change			X	X	X	X	X	X	—
Tensile strength				X	X	X	X		D882 D638 D412
Tear strength			X			X			D1004 D751
Hardness				X	X		X		D785 D2240
Compression set				X					D395

Table 8. Product compatibility summary

Material class	Product	Compatibility rating
Adhesives	Epon 901/B-3	C
	Epon 828/Z	NC
	HT 424	—
Coatings— encapsulants	Solithane 113/300	M
	Stycast 1090/11	C
	Stycast 1095/11	C
	Scotchcast 281 A/B	C
Tapes	Mystik 7020	M
	Mystik 7503	M
Films	FEP 200A	C
	Tedlar 100 BG 30 WH	M
	Mylar 100A	M
	Kapton 200 X H667	C
Hardware and structural materials	Micarta H 17511	C
	Micarta 20201-2	C
	Micarta 65M25	C
Elastomers	Viton B	M
	1050-70 (silicone)	NC
Coated fabrics	SRG 1810	C
	Armalon 98-101	C
	Pyre ML—Glass fabric	NC

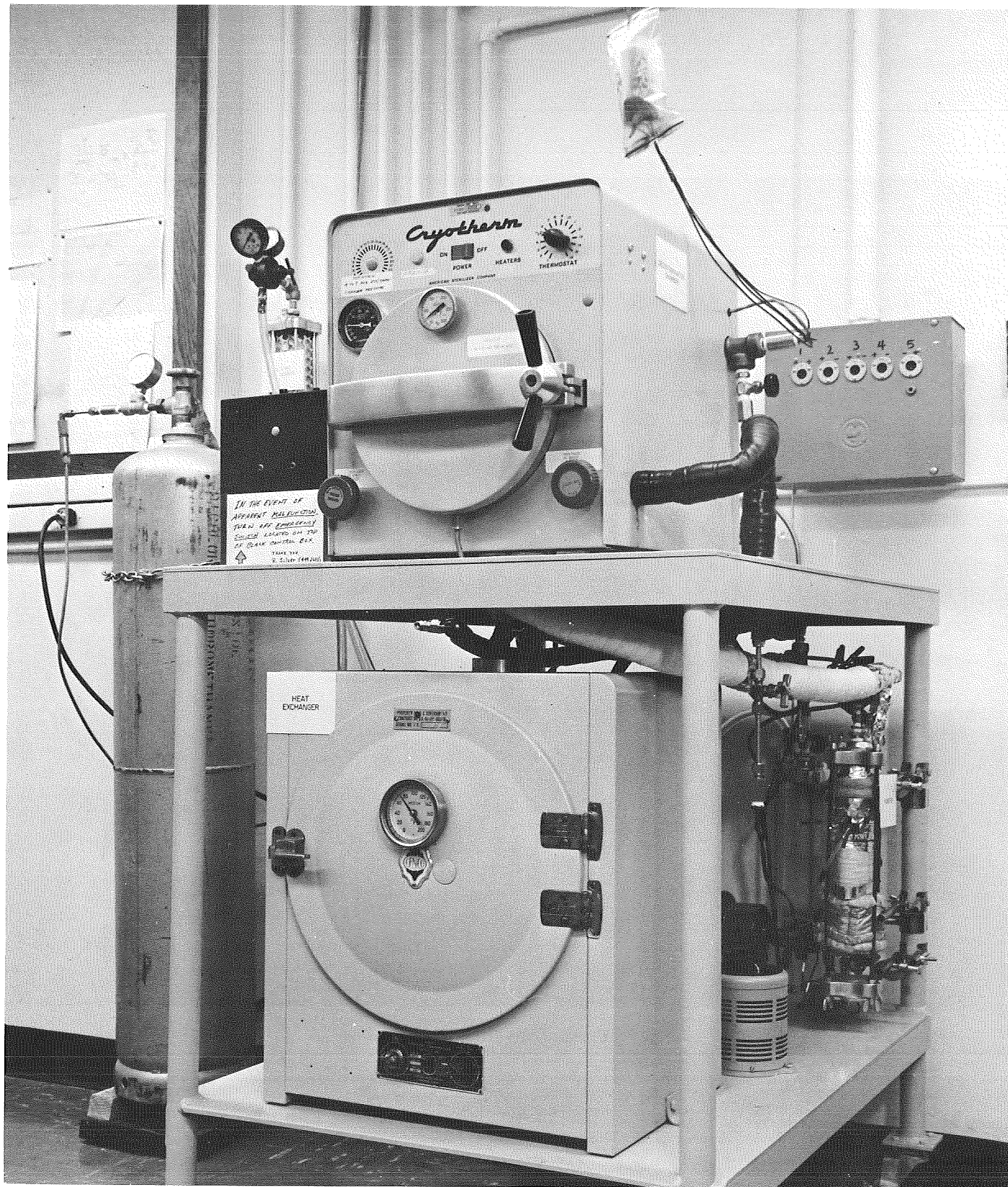


Fig. 7. Ethylene oxide decontamination apparatus

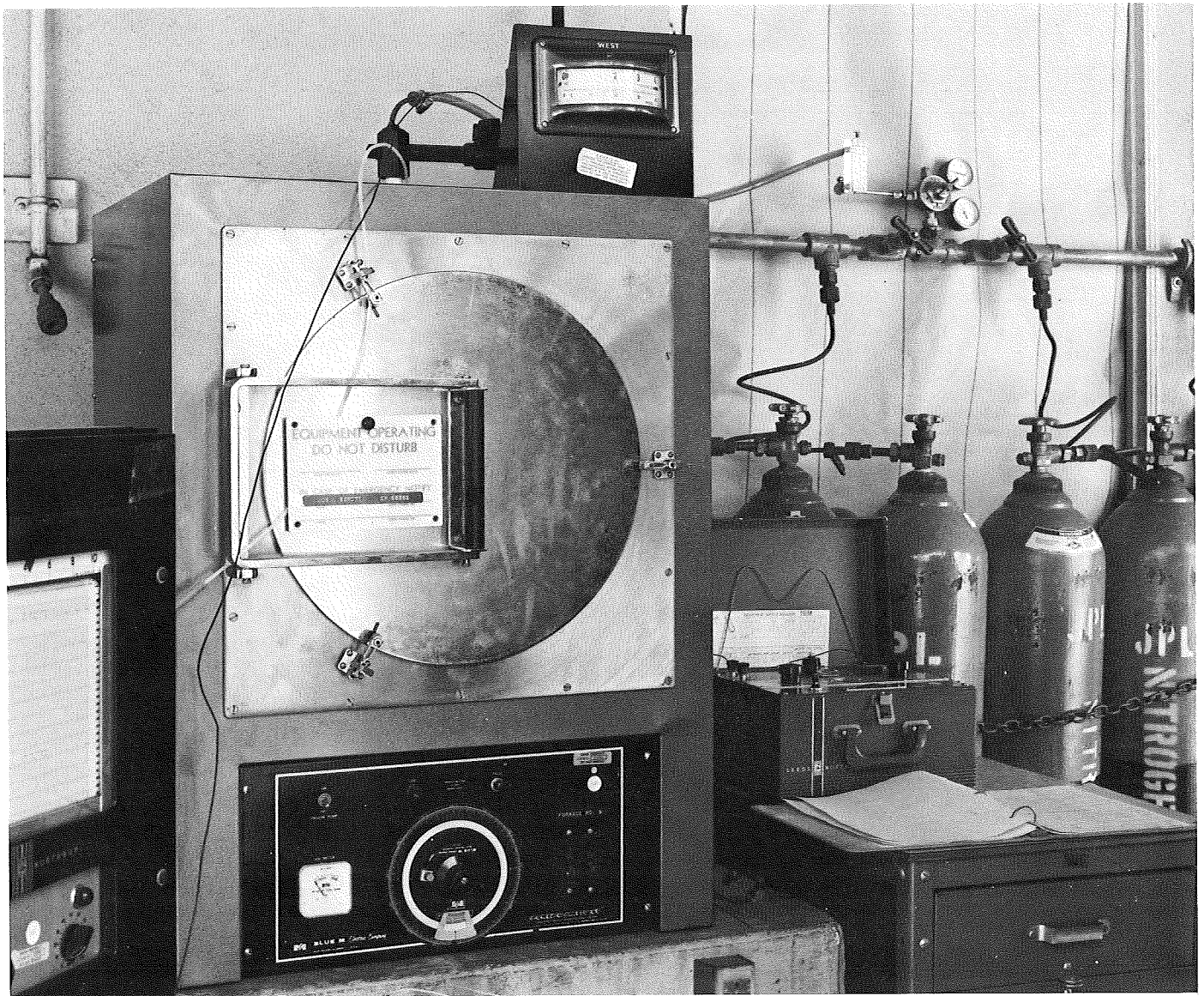


Fig. 8. Thermal-sterilization apparatus

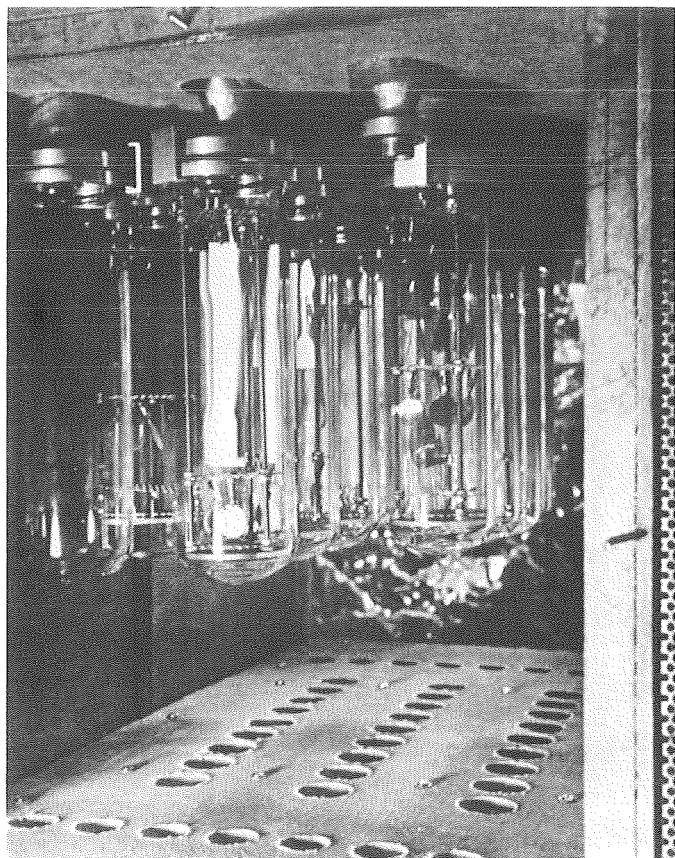


Fig. 9. Thermal-vacuum exposure apparatus

chemical interactions. Incompatibility may arise from low molecular weight polymer migration or exchange between the contacting materials (e.g., silicone polymer migration into structural adhesives or structural thermoplastics). Alkali-induced hydrolysis may be another possible incompatibility. This may be exemplified by possible polyester or polyimide degradation as result

of contact with materials containing residual alkali (phenolic resins).

The selected material combinations are currently undergoing various environmental exposures. The studies within this task are scheduled for completion at the end of Fiscal Year 1970.

References

1. Rydelek, R. F., and Landis, A. L., *Study of the Effects of Ethylene Oxide-Freon 12 Upon Properties of Polymers and Metallic Surfaces*, Final Report, JPL Contract 951003. Hughes Aircraft Company, Culver City, Calif., June 1966.
2. Kalfayan, S. H., and Campbell, B. A., *Effects of the Thermal Sterilization Procedure on Polymeric Products*, Technical Report 32-973. Jet Propulsion Laboratory, Pasadena, Calif., Nov. 15, 1966.
3. Kalfayan, S. H., Campbell, B. A., and Silver, R. H., *Effects of Ethylene Oxide-Freon 12 Decontamination and Dry Heat Sterilization Procedures on Polymeric Products*, Technical Report 32-1160. Jet Propulsion Laboratory, Pasadena, Calif., Sept. 15, 1967.
4. Muraca, R. F., and Whittick, J. S., *Polymers for Spacecraft Applications*, Final Report, JPL Contract 950745. Stanford Research Institute, Menlo Park, Calif., Sept. 1967.
5. Lee, S. M., and Licari, J. J., *Effects of Decontamination and Sterilization on Spacecraft Polymeric Materials*, Final Report, JPL Contract 951566. Autonetics, Division of North American Rockwell Corp., Anaheim, Calif., Jan. 1968.
6. Muraca, R. F., *Polymers for Spacecraft Hardware: Test Equipment Design*, Interim Report No. 3, Part II, JPL Contract 950745. Stanford Research Institute, Menlo Park, Calif., Dec. 1966.
7. Roper, W. D., *Effects of Decontamination, Sterilization, and Thermal Vacuum on Spacecraft Polymeric Products*, Technical Report 32-1411. Jet Propulsion Laboratory, Pasadena, Calif., June 30, 1969.

XIV. Applied Mechanics

ENGINEERING MECHANICS DIVISION

A. High-Impact Dynamic-Response Analysis of Nonlinear Structures, K. K. Gupta

1. Statement of the Problem

The development of electronic equipment capable of hard-landing survival is of primary interest in connection with the unmanned exploration of the planets. In particular, considerable effort has been expended towards the development of a battery power supply required for a proposed hard landing on Mars in 1973 (Ref. 1). Such cell structures are expected to be subjected to a shock level of about 4000 *g* (square pulse for 1 ms) from an impact velocity of 120 ft/s, the relevant data being calculated from available information pertaining to entry conditions based on the gravitational and atmospheric conditions on Mars. Under such high impact the battery cell structure and its components, in particular the cell plates, are expected to undergo large deformations, thereby introducing structural nonlinearities into the system.

2. Method of Solution

A vital preliminary for the dynamic-response analysis of linear structures involves the determination of the

eigenvalues and associated modes. Although efficient numerical procedures (Ref. 2)¹ are now available for such solutions, the classical approach becomes somewhat impractical for the solution of structures involving a very large number of degrees of freedom (*r*). For instance, while in the step-by-step method, the solution time is proportional to $r \times N$, *N* being the number of time steps; in the normal mode method, the corresponding factor is r^3 . In such cases and also in particular when the structures are nonlinear in character, it is necessary to adopt an alternative formulation involving a step-by-step incremental procedure. With such incremental procedures it is further possible to deal with nonlinear damping matrices.

The incremental procedure has been adopted earlier by several authors (Refs. 3 and 4) for the solution of nonlinear problems in which the accelerations are usually assumed to vary linearly during each limited time increment. The set of dynamic equations may then be transformed into a convenient pseudostatic form, the solution

¹Also Gupta, K. K., "Vibration of Frames and Other Structures with Banded Stiffness Matrix," *Int. J. Numerical Methods in Engineering* (to be published).

of which yields the incremental results. Thus the procedure involves the repeated solution of a set of equations having the same order as the degrees of freedom of the structure.

In connection with the present work, a step-by-step numerical integration procedure employing the Runge-Kutta extrapolation formula of $O(h^5)$ truncation error, in conjunction with the finite element matrix displacement approach utilizing quadrilateral shell elements, has been adopted for the solution. The generalized computer program developed for such analysis is designed to cater to geometrical nonlinearities caused by large deformations and may further be easily extended to take into account material nonlinearity. The present analysis is essentially a nodewise predictor-corrector method and has the following primary steps within a small time interval $h(n, n+1)$.

Step 1

To compute acceleration at beginning of step:

$$\ddot{\mathbf{q}}^n = \frac{1}{\mathbf{m}^n} [\mathbf{P}^n - \mathbf{S}^n - \mathbf{C}^n \dot{\mathbf{q}}^n] \quad (1)$$

and store

$$\mathbf{K1} = h\ddot{\mathbf{q}}^n$$

where

\mathbf{q} = displacement vector

\mathbf{m} = nodal mass vector/matrix

\mathbf{P} = nodal external loading vector

\mathbf{S} = nodal forces vector

\mathbf{C} = damping matrix

Step 2

To extrapolate values of incremental deformations and velocity at half-time intervals in order to compute corresponding acceleration:

$$\ddot{\mathbf{q}}^{n+1/2} = \frac{1}{\mathbf{m}^n} \left[\mathbf{P}^n - \mathbf{S}^n - \mathbf{K}^n \left(\frac{h}{2} \dot{\mathbf{q}}^n + \frac{h}{8} \mathbf{K1} \right) - \mathbf{C}^n \left(\dot{\mathbf{q}}^n + \frac{\mathbf{K1}}{2} \right) \right] \quad (2)$$

and store

$$\mathbf{K2} = h\ddot{\mathbf{q}}^{n+1/2}$$

where \mathbf{K} is the stiffness matrix.

Step 3

To compute revised acceleration values at half-time step:

$$\ddot{\mathbf{q}}_R^{n+1/2} = \frac{1}{\mathbf{m}^n} \left[\mathbf{P}^n - \mathbf{S}^n - \mathbf{K}^n \left(\frac{h}{2} \dot{\mathbf{q}}^n + \frac{h}{8} \mathbf{K2} \right) - \mathbf{C}^n \left(\dot{\mathbf{q}}^n + \frac{\mathbf{K2}}{2} \right) \right] \quad (3)$$

and store

$$\mathbf{K3} = h\ddot{\mathbf{q}}_R^{n+1/2}$$

Step 4

To form revised stiffness matrix based on \mathbf{q}_R^{n+1} configuration in order to compute acceleration at end of $n+1$ step:

$$\ddot{\mathbf{q}}^{n+1} = \frac{1}{\mathbf{m}^n} \left[\mathbf{P}^n - \mathbf{S}^n - \mathbf{K}_R^{n+1/2} \left(h\dot{\mathbf{q}}^n + \frac{h}{2} \mathbf{K3} \right) - \mathbf{C}^n (\dot{\mathbf{q}}^n + \mathbf{K3}) \right] \quad (4)$$

and store

$$\mathbf{K4} = h\ddot{\mathbf{q}}^{n+1}$$

Step 5

To update deformation, velocity, and nodal forces files:

$$\mathbf{q}^{n+1} = \mathbf{q}^n + h \left[\dot{\mathbf{q}}^n + \frac{1}{6} (\mathbf{K1} + \mathbf{K2} + \mathbf{K3}) \right] = \mathbf{q}^n + \mathbf{qI}^{n+1} \quad (5)$$

$$\dot{\mathbf{q}}^{n+1} = \dot{\mathbf{q}}^n + \frac{1}{6} [\mathbf{K1} + 2\mathbf{K2} + 2\mathbf{K3} + \mathbf{K4}] = \dot{\mathbf{q}}^n + \dot{\mathbf{q}}\mathbf{I}^{n+1} \quad (6)$$

$$\mathbf{S}_{(n+1)}^{n+1} = \mathbf{S}_{(n)}^{n+1} + \mathbf{K}_R^{n+1/2} \mathbf{qI}^{n+1} \quad (7)$$

Curved quadrilateral elements, formed by assembling such triangular elements (Ref. 5), have been utilized to

idealize the cell structures; the element stresses computed during the process relate to the center area of the element.

3. Data Used for Numerical Example

A typical plate (Fig. 1) was analyzed on a UNIVAC 1108 using the present program. The results are presented in Figs. 2-5. The primary relevant data were as follows:

Impact velocity, ft/s	120
Time increment used for the step-by-step method, μ s	5.20
Number of increments	24
Number of nodes	55
Number of elements	40
Young's modulus E , 10^5 lb/in. ²	3.75
Poisson's ratio μ	0.3 (assumed 0.0 for this example)

Detailed results on the solution of the cell casing will be presented in a later article.

4. Conclusions

Important conclusions and advantages of the present method over other known procedures are summarized below.

- (1) The present analysis is stable, accurate and extremely fast computation-wise.

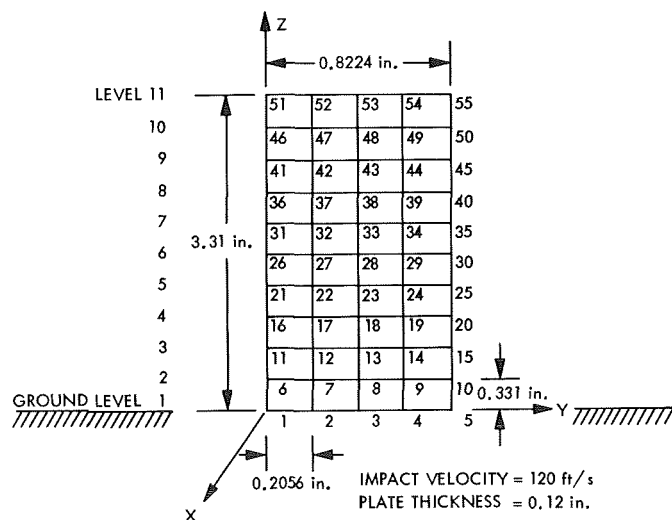


Fig. 1. Typical plate with finite element idealization

- (2) For short impact contact time, the process appears to be very stable and its use is justified over the usual method involving repeated solution of simultaneous equations.

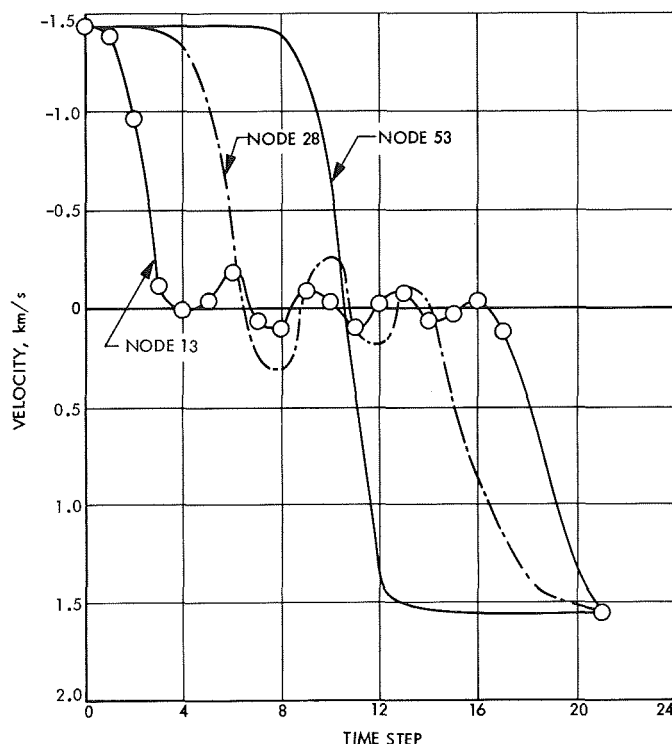


Fig. 2. Propagation of velocity in typical nodes of a panel subjected to high impact (model Fig. 1)

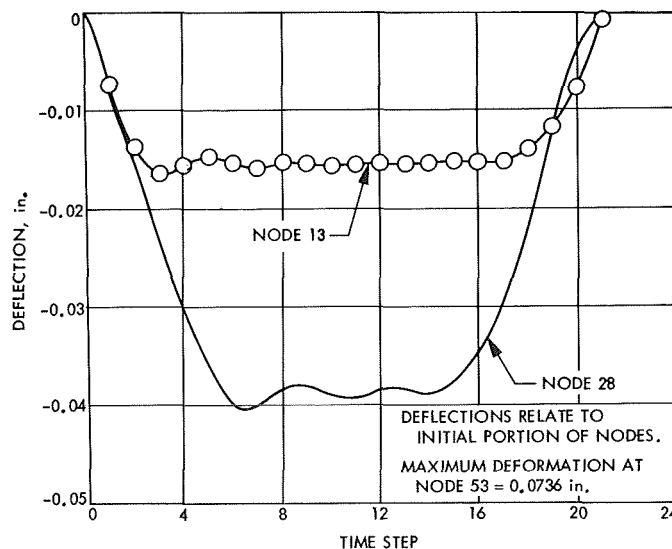


Fig. 3. Propagation of typical nodal deformations during impact

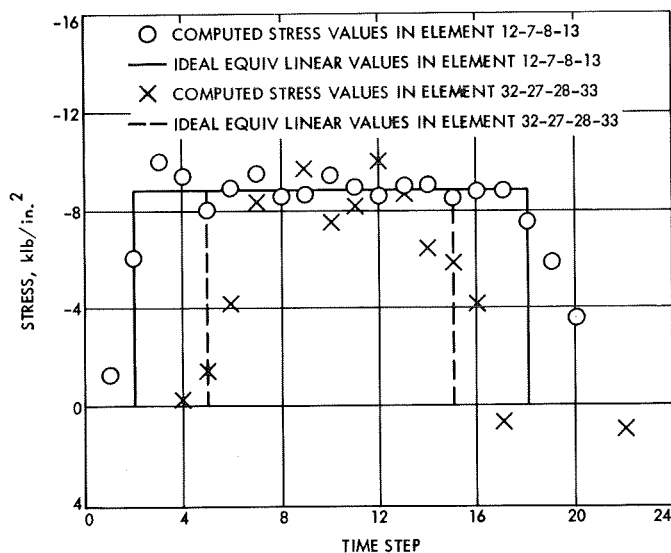


Fig. 4. Stress propagation in typical panel elements during impact (model Fig. 1)

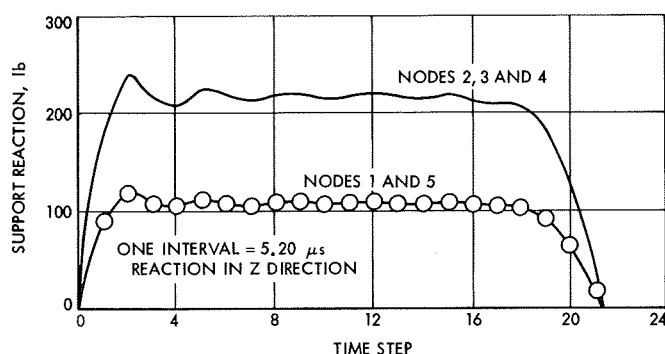


Fig. 5. Propagation of reactions at nodes in contact with support, throughout impact period and until "lift-off" of structure (model Fig. 1)

- (3) Much less computer storage is necessary for the present case when compared with the usual method. Since computations are performed node-wise, the storage required for the stiffness matrix has the dimension $(6, m+1)$, m being the number of nodes connected to any particular joint. In fact, about 10 times the number of degrees of freedom is the approximate maximum storage requirement.
- (4) There is a considerable reduction in solution time involving only repeated simple computations.
- (5) No strict node numbering is necessary.
- (6) The analysis is independent of the form of both the stiffness \mathbf{K} and damping \mathbf{C} matrices, which may indeed be nonlinear in character.

- (7) For very small impact contact time, as the case often is for high-impact problems, the damping appears to have little effect on the overall solution.

References

1. Heat Sterilizable and Impact Resistant Ni-Cd Battery Development, Modification No. 6, JPL Contract 951972, Texas Instruments, Inc., Dallas, Texas, June 1969.
2. Gupta, K. K., "Free Vibrations of Single-Branch Structural Systems," *J. Inst. Math. Appl.*, Vol. 5, No. 3, Sept. 1969.
3. Clough, R. W., Bensuka, K. L., and Wilson, E. L., "Inelastic Earthquake Response of Tall Buildings," World Conference on Earthquake Engineering, Wellington, New Zealand, 1965.
4. Chan, S. P., Cox, H. L., and Benfield, W. A., "Transient Analysis of Forced Vibrations of Complex Structural-Mechanical Systems," *J. Roy. Aeronaut. Soc.*, July 1962.
5. Utku, S., "Stiffness Matrices for Thin Triangular Elements of Non-zero Gaussian Curvature," *AIAA J.*, Vol. 5, No. 9, pp. 1650-1667, Sept. 1967.

B. Eigenvalue of $(\mathbf{B} - \lambda \mathbf{A}^*) \mathbf{y} = 0$ With Positive Definite Band Symmetric \mathbf{B} and Band Hermitian \mathbf{A}^* and Its Application to Natural Frequency Analysis of Flexible Space Vehicles, K. K. Gupta²

1. Introduction

Modern spacecraft and artificial satellites have certain common dynamic and physical characteristics, although they occur in widely varying configurations. Thus current spacecraft are often composed of one or more rigid bodies, each with one or more flexible deployable appendages. Also during the orbiting motion of the spacecraft such appendages as large antennas, solar cell arrays or propulsion devices change their relative orientation, thereby introducing certain additional forces in the usual equilibrium equation.

Among the three distinct analysis methods for flexible spacecraft, namely the discrete coordinate formulation, the hybrid coordinate formulation, and the vehicle normal coordinate formulation, the second method is the one most widely used. The method is based on the possibility of separating a given spacecraft into a number of structural subsystems. Any such subsystem may either be a linearly elastic flexible appendage, undergoing small deformation, or a rigid body (Ref. 1).

²The author gratefully acknowledges the assistance of Dr. E. W. Heer, Supervisor of the Structures and Dynamics Research Group.

The equilibrium equation of a flexible appendage may be derived with reference to Fig. 6, in which A is a flexible elastic structure attached to a rigid base B undergoing arbitrary angular motion of velocity ω . In addition, the appendage A may sometimes have an angular velocity Ω^a relative to the rigid body B , which of course greatly complicates the problem by introducing nonlinearity.

2. Free Vibrations of Rotating Structures

Such equation of equilibrium for a flexible appendage, idealized as a discrete structural system, will have the following matrix form:

$$\mathbf{M}'\ddot{\mathbf{q}} + \mathbf{E}'\dot{\mathbf{q}} + \mathbf{C}'\dot{\mathbf{q}} + \mathbf{K}'\mathbf{q} + \mathbf{A}'\mathbf{q} = \mathbf{L}' \quad (1)$$

where

\mathbf{M}' = modified mass matrix

\mathbf{E}' = damping matrix

\mathbf{K}' = modified stiffness matrix which includes centrifugal force effect

\mathbf{C}' = matrix which includes coriolis force effect

\mathbf{A}' = matrix incorporating effect of angular acceleration, if any

\mathbf{L}' = forcing function vector

\mathbf{q} = deformation vector

and in which the matrices also incorporate the effect of shift in the center of gravity of masses. The effect of

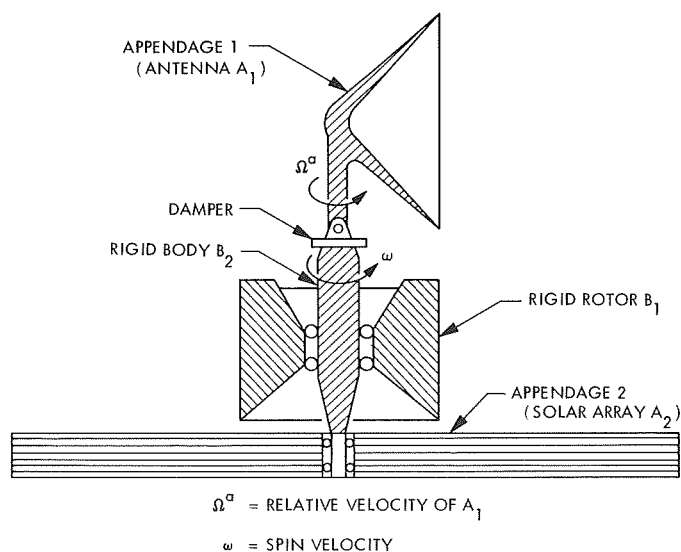


Fig. 6. Typical tri-spin satellite showing arrangement of flexible appendages and rigid bodies

center-of-gravity shift however is almost insignificant being proportional to nodal mass/ μ , where μ is the total mass of the vehicle, and hence may be neglected. Considering now the most commonly occurring case when only the rigid base is undergoing a uniform angular motion, the equation of free motion may be derived by reducing Eq. (1) to the form

$$\mathbf{M}\ddot{\mathbf{q}} + \mathbf{C}\dot{\mathbf{q}} + \mathbf{K}\mathbf{q} = 0 \quad (2)$$

in which \mathbf{M} is the original banded symmetric distributed mass matrix, \mathbf{C} the modified tridiagonal skew symmetric Coriolis force matrix, and \mathbf{K} the further modified symmetric stiffness matrix incorporating the centrifugal force effect, in which the original bandwidth has been retained. Similar equations are also obtained in connection with the determination of natural frequencies of flexible gravity-gradient satellites (Ref. 2).

The solution of Eq. (2) may be achieved by representing it in the following form:

$$\begin{bmatrix} \mathbf{0} & -\mathbf{M} \\ \mathbf{M} & \mathbf{C} \end{bmatrix} \begin{Bmatrix} \ddot{\mathbf{q}} \\ \dot{\mathbf{q}} \end{Bmatrix} + \begin{bmatrix} \mathbf{M} & \mathbf{0} \\ \mathbf{0} & \mathbf{K} \end{bmatrix} \begin{Bmatrix} \dot{\mathbf{q}} \\ \mathbf{q} \end{Bmatrix} = \{0\} \quad (3)$$

which may be written as

$$\mathbf{A}\dot{\mathbf{y}} + \mathbf{B}\mathbf{y} = 0 \quad (4)$$

where \mathbf{A} is skew symmetric, \mathbf{B} being symmetric and positive definite, and $\mathbf{y} = \dot{\mathbf{q}}/\mathbf{q}$. Solution of Eq. (4) may be taken as $\mathbf{y} = e^{pt}$ (Ref. 3, p. 319), and substituting the same in Eq. (4) yields

$$(\mathbf{B} + p\mathbf{A})\mathbf{y} = 0 \quad (5)$$

which may finally be reduced to the eigenvalue problem

$$(\mathbf{A}^* - \omega\mathbf{B})\mathbf{y} = 0 \quad (6)$$

in which $\mathbf{A}^* = i^*\mathbf{A}$ is a Hermitian matrix, the roots $\omega = 1/i^*p$ being real and nontrivial solution of

$$|\mathbf{A}^* - \omega\mathbf{B}| = 0 \quad (7)$$

i^* being $(-1)^{1/2}$.

The eigenvalues of Eq. (7) are real and will exist in pairs $\omega_1, -\omega_1; \omega_2, -\omega_2; \dots; \omega_n, -\omega_n$, whereas the roots corresponding to the original form of Eq. (5) will be pure

imaginary being simply $1/i^*\omega$ and occurring in complex conjugate pairs. Associated eigenvectors for both cases are the same and will occur as pairs of complex conjugate vectors. Further, Eq. (7) is valid for any stable structure and hence also valid for any spacecraft in its entirety, being composed of various subsystems of appendices and rigid bases.

A great majority of structures, when idealized as a discrete structural system by, say, the finite element method, yields highly banded \mathbf{K} and \mathbf{M} matrices. For realistic and accurate analysis of such structures, it is often necessary to choose finer mesh sizes of the finite elements, which on the other hand increases the number of degrees of freedom of the structural system and hence the order of the \mathbf{K} and \mathbf{M} matrices. It is then extremely important to retain such banded form during all subsequent analysis, both from the point of view of economy of computer storage and also run time.

The solution of Eq. (7) may be obtained by presently known standard procedures. Thus, Eq. (6) may be re-written in the equivalent form

$$(\mathbf{D}^* - \omega \mathbf{I})\mathbf{u} = 0 \quad (8)$$

where

$$\mathbf{D}^* = \mathbf{L}^{-1} \mathbf{A}^* (\mathbf{L}^T)^{-1}, \quad \mathbf{u} = \mathbf{L}^T \mathbf{y} \quad (9a)$$

\mathbf{I} being the unit matrix and \mathbf{L} a lower triangular matrix obtained from the Choleski decomposition

$$\mathbf{B} = \mathbf{L} \mathbf{L}^T \quad (9b)$$

which can be carried out since both \mathbf{K} and \mathbf{M} matrices are positive definite.

Equation (9) is the standard eigenvalue problem for the Hermitian matrix \mathbf{D}^* and may be solved by, say, the method of Givens, Householder, or Jacobi (Ref. 4). However, this procedure is extremely inefficient and almost prohibitive excepting for a few trivial simple structural problems, since \mathbf{D}^* is a full matrix.

An alternative procedure is described in the later sections which exploit the band form, however, and which utilize the \mathbf{K} , \mathbf{M} , and \mathbf{C} matrices of Eq. (3) in their basic form and determines the eigenvalues and associated eigenvectors directly, employing only a small amount of working storage in addition to that required to store the elements in the symmetric half of \mathbf{K} , \mathbf{M} , and \mathbf{C} . This is

achieved by first establishing the validity of the Sturm sequence property for the case of Eq. (7) and then by evolving an efficient and numerically stable algorithm for the determination of the eigenvalues based on the Sturm sequence method. The eigenvectors are finally computed by the inverse iteration technique.

3. The Sturm Sequence Method for the Hermitian Case

For any real symmetric matrix \mathbf{R} the leading principal minors of $\mathbf{R} - \omega \mathbf{I}$ is known to possess the Sturm sequence property. Thus, for a given value ω , the number of changes (agreements) in sign of consecutive members of the sequence $p_r(\omega)$, starting with $p_0(\omega) = 1$, is equal to the number of eigenvalues of $\mathbf{R} - \omega \mathbf{I}$ smaller (greater) than ω in algebraic value. This property enables us to compute any individual eigenvalue, the r th say, without computing any other. Further, in connection with the usual structural eigenvalue problems, when \mathbf{M} is symmetric and \mathbf{K} is symmetric and positive definite, the leading principal minors $p_r(\omega)$ of $\mathbf{M} - \omega \mathbf{K}$ also form a Sturm sequence, a property which considerably simplifies solution of such problems (Ref. 5). The advantages of utilizing an eigenvalue routine based on the Sturm sequence method, for the solution of practical structural problems, have been amply demonstrated elsewhere.³

It may further be derived that for any Hermitian matrix \mathbf{D}^* the sequence $p_r(\omega)$ of $\mathbf{D}^* - \omega \mathbf{I}$ is a Sturm sequence. Also the validity of the Sturm sequence property for Eq. (7) has been recently established.⁴ However, from a computational point of view, it would be advantageous to work with an alternative form of Eq. (7) as below:

$$|\mathbf{B} - \lambda \mathbf{A}^*| = 0 \quad (10)$$

in which $\lambda = 1/\omega$, and it is quite straightforward to deduce that $p_r(\lambda)$ of Eq. (10) also forms a Sturm sequence since it is an equivalent form of Eq. (7).

4. Basic Numerical Schemes

The Sturm sequence method enables computation of all roots of $|\mathbf{B} - \lambda \mathbf{A}^*|$ or any particular ones as desired.

³Gupta, K. K., "Vibration of Frames and Other Structures with Banded Stiffness Matrix," to be published in *Int. J. Numerical Methods in Engineering*.

⁴Gupta, K. K., "Analysis of Free Vibration of Spacecraft and Other Spinning Structures with Stiffness Matrix of Banded Configuration," to be submitted to the *Int. J. Numerical Methods in Engineering* for possible publication.

Thus, once the range of values for λ is specified, the associated computer program determines the number of eigenvalues lying in that range and then any particular desired root is isolated by repeated bisection. Once a root has been isolated, it is then accurately located with super-linear convergence by a procedure based on the *regula falsi* method (Footnote 3) involving a combination of bisection and successive linear interpolation. An alternative method involving successive linear interpolation between weighted function values of opposite sign has also proved to be very satisfactory (Ref. 5).

A vital part of the eigenvalue scheme is the determination of the leading principal minors and this cannot be obtained in general by employing gaussian elimination with partial or complete pivoting (Ref. 4, p. 239). But for numerical stability, some form of pivoting is essential which, however, must not destroy the information necessary to evaluate the signs of the required determinants, and which ensures that multipliers do not exceed unity in modulus, thus preserving the numerical stability. This is achieved by developing a special variant of the gaussian elimination, which yields the principal minors during the reduction procedure. In this process, the $(r+1)$ th row is not involved until the r th major step and the $(r+1)$ th principal minor $f_{r+1}(\lambda)$ is immediately obtained at the end of the current $(r+1)$ th step.

In connection with the present problem of Eq. (10), a numerically stable algorithm has been developed which also takes into account the special symmetric configuration of the matrices of Eq. (10) (Fig. 7), thereby reducing the computer storage requirement to a minimum (Footnote 4). Thus, in addition to storing one-half of the elements of \mathbf{M} , \mathbf{K} , and \mathbf{C} matrices, the procedure needs only

two modest working stores \mathbf{D} and $\mathbf{D}\mathbf{I}$, both of dimension $(m+1)(2m+1)$, $2m+1$ being the bandwidth.

5. Inverse Iteration Scheme for Eigenvectors

Referring to Eq. (8), the eigenvector \mathbf{u} corresponding to \mathbf{D}^* can be computed by the inverse iteration scheme (Ref. 4, p. 321).

$$(\mathbf{D}^* - \omega \mathbf{I}) \mathbf{u}_{r+1} = \mathbf{u}_r \quad (11)$$

where \mathbf{u}_r may be assigned arbitrarily. Unless \mathbf{u}_r is pathologically deficient in \mathbf{u} , \mathbf{u}_{r+1} tends to \mathbf{u} and, at most, two iterations should be sufficient to achieve working accuracy. Returning now to the present characteristic equation $(\mathbf{B} - \lambda \mathbf{A}^*) \mathbf{y} = 0$, the eigenvectors may be computed by first forming the iteration scheme. Thus, Eq. (11) may be rewritten as

$$(\mathbf{L}^{-1} \mathbf{A}^* (\mathbf{L}^T)^{-1} - \omega \mathbf{I}) \mathbf{L}^T \mathbf{y}_{r+1} = \mathbf{L}^T \mathbf{y}_r \quad (12)$$

multiplying both sides of Eq. (12) by \mathbf{L} and noting that $\mathbf{B} = \mathbf{L}\mathbf{L}^T$, one obtains

$$(\mathbf{A}^* - \omega \mathbf{B}) \mathbf{y}_{r+1} = \mathbf{B} \mathbf{y}_r \quad (13)$$

which may finally be reduced to the required iteration scheme

$$(\mathbf{B} - \lambda \mathbf{A}^*) \mathbf{y}_{r+1} = -\frac{1}{\lambda} \mathbf{B} \mathbf{y}_r \quad (14)$$

Equation (14) may be solved by using the same special variant of gaussian elimination used for the Sturm sequence procedure and, at most, two iterations are sufficient

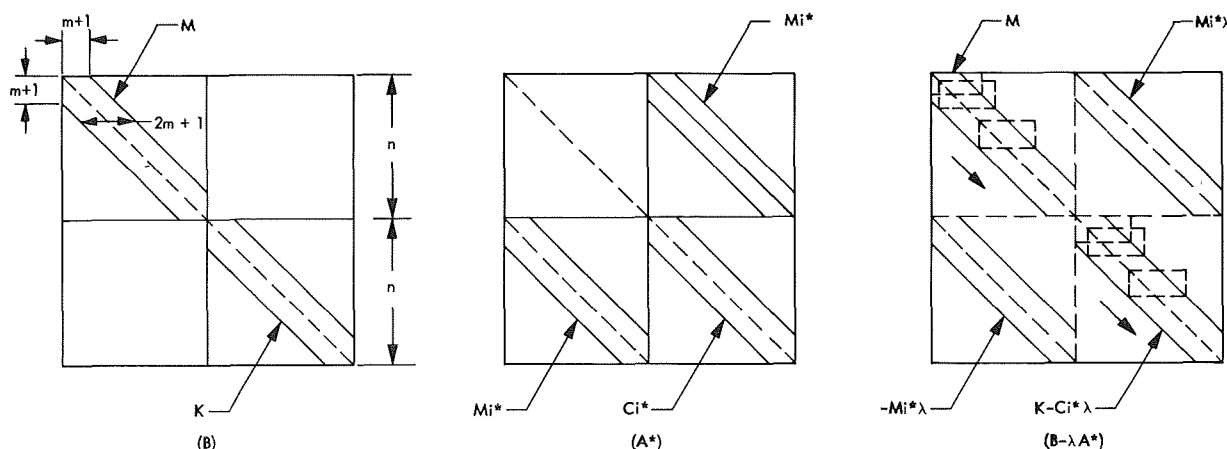


Fig. 7. Arrangement of matrices for $|\mathbf{B} - \lambda \mathbf{A}^*|$ formulation

to yield the eigenvectors, unless \mathbf{y}_r is pathologically deficient in \mathbf{y} . It may be noted that the eigenvector corresponding to equation $(\mathbf{B} - \lambda\mathbf{A}^*)\mathbf{y}$ is the same as that of Eq. (5).

6. Conclusions

The following important conclusions may be drawn in connection with the analysis presented in this article:

(1) The validity of Sturm sequence property for the $\mathbf{B} - \lambda\mathbf{A}^*$ case, \mathbf{B} being banded real symmetric and positive definite whereas \mathbf{A}^* is banded Hermitian, has now been established.

(2) Solution of related structural eigenvalue problem is considerably simplified by adopting the Sturm sequence method. Moreover the computer programming is simplified.

(3) Significant savings in computer storage is achieved by using the present method since it exploits the banded form of the related matrices. Thus, it is only necessary to form and store one-half the elements of the highly banded \mathbf{M} , \mathbf{K} , and \mathbf{C} matrices; these are then used directly to determine the eigenvalues and associated modes. In addition to the usual storage, an additional modest working space of $2(m+1)(2m+1)$ elements only is necessary to find the eigenvalues. However, the usual procedures involve the solution of Eq. (8) requiring storage of $(2n \times n)$ elements for the \mathbf{D}^* matrix in addition to the usual storage required for \mathbf{M} , \mathbf{K} , and \mathbf{C} matrices.

(4) Considerable savings in matrix operations, and hence computation time, results in using the Sturm sequence method. Thus, whereas the eigenvalue computation time employing any other standard eigenvalue routine is proportional to $(2n)^3$, the equivalent time required for the present method is proportional to $2nm^2$ only. On the other hand, if the masses are assumed to be lumped, the computation time will be proportional to nm^2 and requiring a work store of only $(m+1)(2m+1)$ elements. This observation is particularly true if only the first few roots are wanted, which however is often the requirement for most practical structures.

(5) An alternative and widely used approach for the solution of Hermitian matrices (Ref. 4, p. 342) involving real number computations requires further doubling of storage ($4n \times 4n$ in this case) and twice the time required for the solution of $\mathbf{D}^* = \lambda\mathbf{I}$ case involving complex number operations. Thus, the alternative formulation compares even less favorably with the Sturm sequence procedure developed herein.

(6) It may further be observed that the larger the structure, i.e., the larger the value of n , the more advantageous it would be to employ the present method over the usual ones, both in terms of computer storage and computation time. Further, the bounds on the structure, both from the point of view of storage and computation time that could be solved, imposed by limitations of the present eigenvalue routines, are now greatly relaxed with the development of the procedure described in this study.

References

1. Likins, P. W., *Dynamics and Control of Flexible Space Vehicles*, Technical Report 32-1329. Jet Propulsion Laboratory, Pasadena, Calif., Feb. 15, 1969.
2. Newton, J. K., and Farrell, J. L., "Natural Frequencies of a Flexible Gravity Gradient Satellite," *J. Spacecraft Rockets*, Vol. 5, No. 5, pp. 560-569, 1968.
3. Hurty, W. C., and Rubinstein, M. F., *Dynamics of Structures*. Prentice-Hall, Inc., Englewood Cliffs, New Jersey, 1964.
4. Wilkinson, J. H., *The Algebraic Eigenvalue Problem*. Clarendon Press, Oxford, London, 1965.
5. Gupta, K. K., "Free Vibrations of Single-Branch Structural Systems," *J. Inst. Math. Appl.*, Vol. 5, No. 3, Sept. 1969.

C. Nonstationary Peak Distribution, M. Shinozuka⁵ and J. N. Yang⁶

1. Introduction

The present study develops a method by which one can establish the distribution function of peak values of a nonstationary narrow-band process within a specified time interval. The narrow-band characteristics are assumed since this study is concerned with the structural response process in relation to the fatigue and the maximum response analysis.

The study of nonstationary peak distribution is of considerable practical importance as well as of theoretical interest since the distribution has not been established in spite of the fact that response processes of a space vehicle are in reality nonstationary. Only work in this area is that by Roberts (Ref. 1) concentrating on the estimation of the expected fatigue damage due to a nonstationary narrow-band process based on a certain cumulative damage hypothesis.

⁵Consultant, Jet Propulsion Laboratory, Pasadena, California, and Professor of Civil Engineering, Department of Civil Engineering and Engineering Mechanics, Columbia University, New York, New York.

⁶NRC-NASA Resident Research Associate, Jet Propulsion Laboratory.

The present study emphasizes the following points unique to the nonstationary problem, the discussion of which was bypassed by Roberts:

- (1) The frequency interpretation of probability involved; this should be made clear because of lack of ergodicity of nonstationary process.
- (2) Analytical form of the peak distribution; for example, assuming that the process is gaussian, how far does the form deviate from the Rayleigh distribution which is the peak distribution of a narrow-band stationary gaussian process.
- (3) The effect of nonstationarity on the maximum peak response within a specified time interval; this is directly related to the gust response factor often used in the structural response analysis to wind load.

2. Results

(1) The nonstationary peak distribution $F_\eta(x;T)$ in the time interval $[0, T]$ is, under the assumption that the process is narrow band and that the apparent period T_0 of the process is much smaller than T ,

$$F_\eta(x;T) = 1 - \frac{\int_0^T \nu^+(x,t) dt}{\int_0^T \nu^+(0,t) dt} \quad (1)$$

where $\nu^+(x,t)$ is the expected rate of upcrossing at level x by the process at time t . The apparent period here is defined as equivalent to the resonance period of the structure since the excitation is assumed to be a wide-band process. It is important to realize that the distribution function $F_\eta(x;T)$ has the following frequency interpretation:

$$F_\eta(x;T) = 1 - \lim_{n \rightarrow \infty} \left\{ \frac{\sum_{k=1}^n k n_x(T)}{\sum_{k=1}^n k n_0(T)} \right\} \quad (2)$$

where $k n_x(T)$ = number of upcrossings at level x by the k th member function of the process, and $k n_0(T)$ = number of upward zero crossings by the same member function both in the time interval $[0, T]$. Equation (2) is an extension of the expression for the peak distribution function widely used for ergodic processes;

$$F_\eta(x; \infty) = 1 - k n_x(\infty) / k n_0(\infty) = 1 - \nu_x^+ / \nu_0^+$$

where ν_x^+ and ν_0^+ are the expected rate of upcrossings at levels x and zero by the ergodic process. Equation (2) indicates that if a structure is subjected to a repeated application of a nonstationary excitation (independently each time), the response peak distribution approaches the form given by Eq. (1) as the number of applications increases. Also, in view of the lack of alternative methods at this time, it appears reasonable to use Eq. (1) for the peak distribution within the same member function for approximation.

(2) Assuming that the process is gaussian and evaluating Eq. (1) numerically with T as a parameter, $F_\eta(x;T)$ is shown to be the Weibull distribution:

$$F_\eta(x;T) = 1 - \exp \{ - \alpha^{-1} (x/\sigma)^\alpha \} \quad (3)$$

where both σ and α depend on the characteristics of the nonstationary process as well as T . The value of α is no longer equal to 2 as in the stationary case. Once σ and α are found, the expected relative fatigue damage $E[\Delta_T]$, as defined by the Palmgren-Miner rule, due to the excitation in the time interval $[0, T]$, is obtained as

$$E[\Delta_T] = N(C_1 \sigma)^b \alpha^{b/\alpha} \Gamma(1 + b/\alpha) / C \quad (4)$$

with

$$N = \int_0^T \nu^+(0,t) dt \simeq T/T_0 \quad (5)$$

where a peak stress S is assumed to be proportional to a peak displacement ($S = C_1 \eta$), C and b are such that the S - N relationship of the fatigue process is given by

$$S^b N = C \quad (6)$$

Equation (4) reduces to the form (Ref. 2) associated with the stationary excitation as α approaches 2 since then $N = \nu^+(0)T$ and $C_1 \sigma = \sigma_s$ = the standard deviation of (stationary) stress response process.

(3) Under the assumption of independence among large peak values, the (asymptotic) distribution function $F_{\eta_m}(x)$ of the largest peak value η_m can be given by

$$F_{\eta_m}(x;T) = \exp \left\{ - \exp \left[- K \alpha^{-1} \left(\frac{x}{\sigma} - K \right) \right] \right\} \quad (7)$$

$$K = (\alpha \ln N)^{1/\alpha} \quad (8)$$

with

$$E[\eta_m] = \{K + 0.5772K^{1-\alpha}\}\sigma \quad (9)$$

$$\sigma_{\eta_m} = 1.28\sigma K^{1-\alpha} \quad (10)$$

where $E[\cdot]$ indicates the expected value, and σ_{η_m} = the standard deviation of η_m . Equations (4-8) again reduce to the well-known corresponding forms associated with the Rayleigh peak distribution when $\alpha = 2.0$.

3. Example

Consider a nonstationary gaussian response process $X(t)$ of a structure to white noise $n(t)$ with $E[n(t)] = 0$ and $E[n(t)n(s)] = D\delta(t-s)$, where $\delta(\cdot)$ is the Dirac delta function. Assume that the structure has the impulse response function $h(t)$ of the form

$$h(t) = H(t) \exp(-\zeta\omega_0 t) \sin \omega t / \omega \quad (11)$$

where ζ = damping ratio, ω_0 = natural frequency, ω = damped natural frequency, and $H(\cdot)$ = Heaviside unit step function.

It then follows that

$$\sigma_x^2 = \sigma_\infty^2 \{q[\zeta^2\omega_0^2 \cos 2\omega t - \zeta\omega_0\omega \sin 2\omega t - \omega_0^2] + 1\} \quad (12)$$

$$\sigma_{\dot{x}}^2 = \sigma_\infty^2 \omega_0^2 \{q[\zeta^2\omega_0^2 \cos 2\omega t + \zeta\omega_0\omega \sin 2\omega t - \omega_0^2] + 1\} \quad (13)$$

$$\rho\sigma_{\dot{x}}\sigma_x = \sigma_\infty^2 \zeta\omega_0^3 q(1 - \cos 2\omega t) \quad (14)$$

in which

$$\sigma_x^2 = E[X^2(t)]$$

$$\sigma_{\dot{x}}^2 = E[\dot{X}^2(t)]$$

$$\rho\sigma_{\dot{x}}\sigma_x = E[X(t)\dot{X}(t)]$$

$$q = \exp[-2\zeta\omega_0 t] / \omega^2$$

and $\sigma_\infty^2 = D/4\zeta\omega_0^3$ is the variance of the corresponding stationary response process.

Because of the assumed normality,

$$\nu^+(x,t) = \frac{1}{2\pi B^{1/2} K_{22}} \exp(-K_{11}x^2/2) \left\{ 1 - x \left(\frac{2\pi}{K_{22}} \right)^{1/2} K_{12} \right. \\ \left. \times \exp\left(\frac{-K_{12}^2 x^2}{2K_{22}} \right) \left[1 - \Phi\left(\frac{K_{12}x}{K_{22}^{1/2}} \right) \right] \right\} \quad (15)$$

where $\Phi(\cdot)$ = standardized normal distribution function, $K_{ij} = i-j$ component of the inverse matrix of the covariance matrix of $X(t)$, and B = the determinant of the covariance matrix:

$$\left. \begin{aligned} B &= \sigma_x^2 \sigma_{\dot{x}}^2 (1 - \rho^2), & K_{11} &= \sigma_x^2 / B \\ K_{12} &= K_{21} = -\rho\sigma_x\sigma_{\dot{x}} / B & K_{22} &= \sigma_{\dot{x}}^2 / B \end{aligned} \right\} \quad (16)$$

In the present example, the peak distribution of the response normalized in terms of σ_∞ with $\omega_0 = 5$ cps, $\zeta = 0.02$ or 0.08 , is considered. Redefining η in the preceding discussion as the peak value of this normalized process, $F_\eta(x;T)$ is computed from Eq. (1) for various values of T and x . Figure 8 plots the result for the case of $T = 3$ s, $\zeta = 0.02$ (heavy solid curve) on the Weibull probability paper. It is practically a straight line indicating that the distribution can be written in the form of Eq. (3) with, in this case, $\alpha = 1.805$ and $\sigma = 0.855$. Also plotted in Fig. 8 is the result of digital simulation; crosses, open circles, and closed circles are based respectively on one, five, and ten simulated sample functions. Chained curve is drawn using the simulation based on 10 sample functions. The observed trend, that these simulated peak distributions become stable as the number n of member functions involved increases, is compatible with the frequency interpretation (Eq. 2). Figure 9 shows the histogram of simulated peaks and troughs together with the theoretical density function associated with Eq. (1). The result of simulation further indicates that the total number of zero crossings in n sample functions can be estimated, in approximation, as $2n\omega_0 T$ (in the present example, $2 \times 10 \times 3 \times 5 = 300$, whereas it is 305 from simulation). The truncated distribution disregarding negative maxima and positive minima can be shown to be on the conservative side in application. Such truncated distribution obtained from simulation is also plotted in Fig. 8 (light solid curve passing through double circles) which is even closer to the solid curve representing Eq. (1). Finally, one can derive the dashed curve in Fig. 8 as an approximation to Eq. (2) by assuming $\rho = 0$. This approximation is referred to as the Rayleigh approximation since it can be interpreted as the weighted

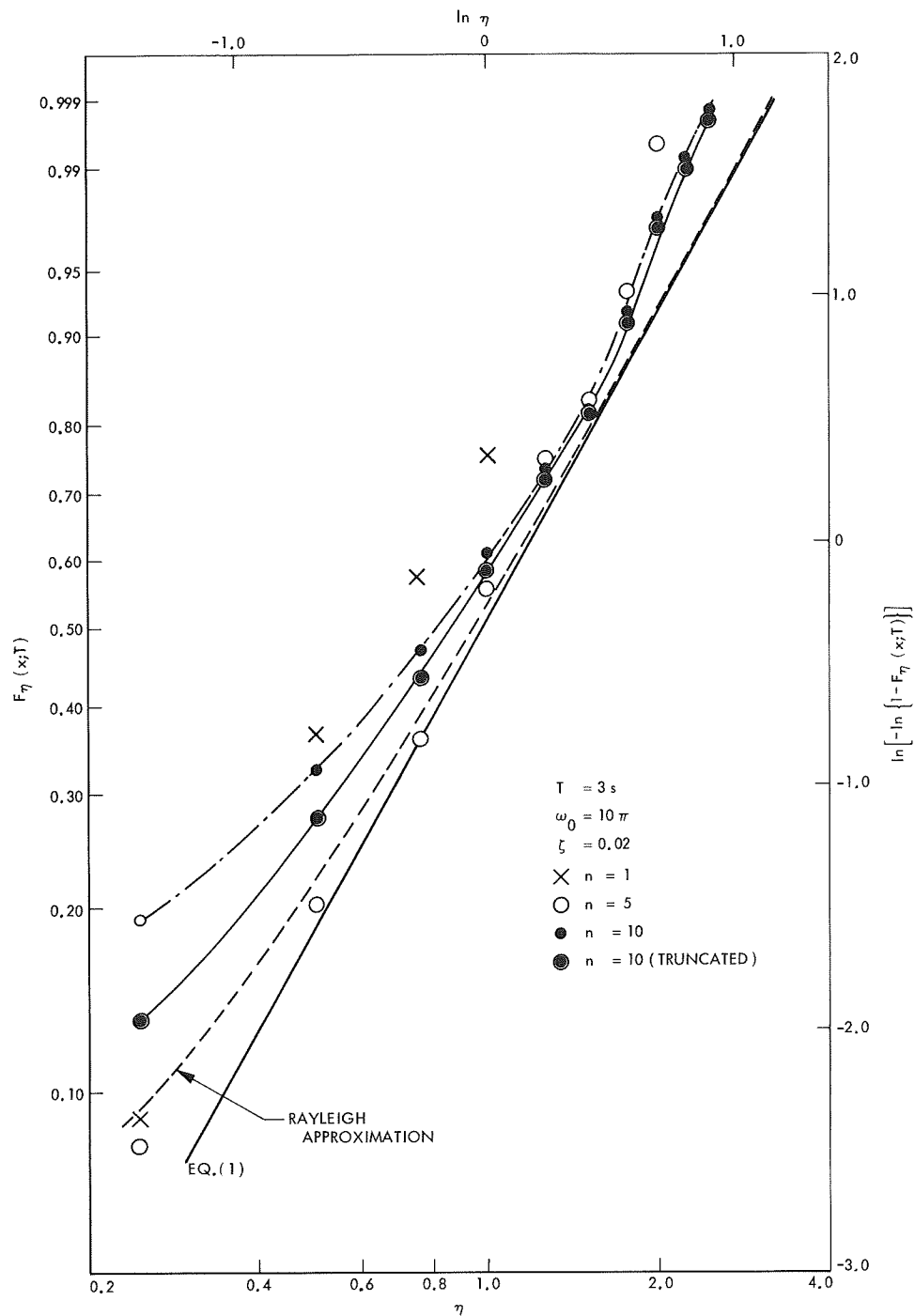


Fig. 8. Distribution function of nonstationary peak

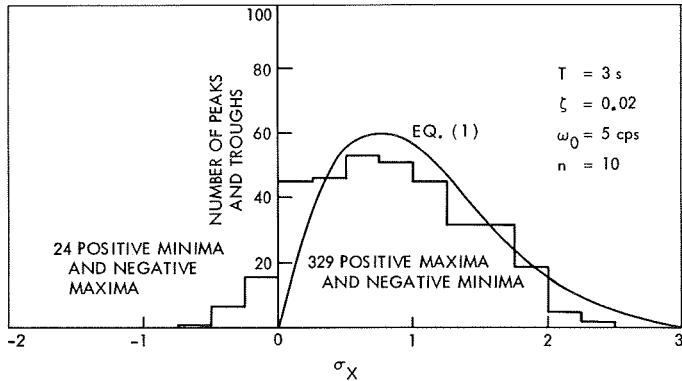


Fig. 9. Histogram of nonstationary peak distribution

average of $1 - \exp(-0.5x^2/\sigma_x^2)$ with respect to the frequency $\sigma_x/(2\pi\sigma_x)$ of occurrence of peaks and troughs along the time.

From the preceding discussion, it appears that Eq. (1) may be used for the nonstationary peak distribution in approximation. Further results show that as ζ decreases and/or T increases, (1) the discrepancy between Eq. (1) and the simulation becomes even less significant, and (2) σ and α (Eq. 3), respectively, approach 1.0 and 2.0 as expected.

References

1. Roberts, J. B., "Structural Fatigue Under Non-Stationary Random Loading," *J. Mech. Eng. Sci.*, Vol. 8, pp. 392-405, 1966.
2. Crandall, S. H., and Mark, W. D., *Random Vibration in Mechanical Systems*, p. 117. Academic Press, New York, 1963.

D. An Analytical and Experimental Study of Heat Transfer in a Simulated Martian Atmosphere,⁷ D. Ting

1. Introduction

During the coming space flights to Mars, thermal design and thermal control will be essential to the success of the mission. The Martian surface lander could conceivably be exposed to ambient temperatures from -180 to 100°F , and wind velocities up to 250 ft/s. The surface atmospheric pressure has been estimated in the neighborhood of 6 mbars and carbon dioxide is believed to be the major constituent present. The low pressure existing on the Martian surface lends itself to speculations as to the validity of continuum flow heat transfer correlations as

well as the manner in which the transport properties vary under such conditions.

Before an analysis can be performed, the above speculations must be clearly defined. It is shown that the experimental results of Johnston, et al (Ref. 1), usually cited to support the argument that the transport properties are functions of pressure below about 0.1 atm (or approximately 100 mbars), can be used as the basis for a tentative flow regime classification. The agreement of this classification scheme with other schemes proposed on direct experimental basis, e.g., those of Schaaf and Chambre (Ref. 2) and Stalder, Goodwin, and Creager (Ref. 3), might serve to indicate that these data have been previously misinterpreted. On this basis it is heuristically concluded that the bulk transport properties remain at values predicted by the kinetic theory while low-pressure phenomenon due to slight rarefaction could be adequately accounted for by classical slip flow analysis.

The flow regime criterion indicates the occurrence of continuum flow over bodies having sizes of practical importance on the Martian surface. Only laminar flow is expected to exist since the Reynolds numbers are small as a result of low gas densities.

Experimental measurements of forced convection heat transfer from simply shaped objects in a simulated Martian atmosphere were performed to further evaluate the reliability of the established solutions and correlations when used under reduced pressures. These experiments were conducted in a low-density wind tunnel at JPL. Average film heat transfer coefficients were measured for a flat plate and a cylinder under various pressures and wind velocities in gas mixtures supposedly present in the Martian atmosphere. The models were electrically heated and fully instrumented so that radiation and conduction losses could be accounted for. The steady-state power input was monitored for computing these coefficients.

Transport property values for calculating the dimensionless parameters were determined using the Chapman-Enskog theory. Constant free stream properties were used for calculating the heat transfer coefficients for the flat plate, while properties evaluated at the film temperature were employed for calculating the overall heat transfer coefficients for the cylinder, as suggested by the originators of the correlations. A maximum deviation of 10% was found between the experimental data and the values computed from the Pohlhausen solution for the flat plate and the correlations for the cylinder.

⁷Work performed by Purdue University under JPL contract.

2. Analytical Results

Average heat transfer coefficients, \bar{h} , for laminar flow over a flat plate and a cylinder were calculated for various Martian atmospheric models.

The mean heat transfer coefficient over a flat plate is given by

$$\bar{h} = 0.664 k Pr^{1/3} \left(\frac{\rho u_{\infty}}{\mu} \right)^{1/2} \frac{1}{L^{1/2}}$$

Denoting the value of \bar{h} for a plate of length L ft by $\bar{h}(L)$, and that for a length of 1 ft by $\bar{h}(1)$, we have

$$\bar{h}(L) = \frac{\bar{h}(1)}{L^{1/2}}$$

where

$$\bar{h}(1) = 0.664 k Pr^{1/3} \left(\frac{\rho u_{\infty}}{\mu} \right)^{1/2}$$

Assuming the Prandtl number for the different atmospheric models to be constant at 0.72, the flat plate relation gives

$$\bar{h}(1) = 0.595 k \left(\frac{\rho u_{\infty}}{\mu} \right)^{1/2}$$

For the cylinder, the local heat transfer coefficient given by Eckert (Ref. 4) for air is

$$\frac{Nu}{Re} = 0.945 - 0.7696 \left(\frac{x}{D} \right)^2 - 0.3478 \left(\frac{x}{D} \right)^4$$

where

D = diameter of cylinder

x = distance above the circumference measured from the stagnation point

Substituting for Nu and Re , we have

$$h_o = k \left(\frac{\rho u_{\infty}}{\mu} \right)^{1/2} \frac{1}{D^{1/2}} f(\theta)$$

where

$$f(\theta) = 0.945 - 0.7696 \left(\frac{x}{D} \right)^2 - 0.3478 \left(\frac{x}{D} \right)^4$$

$$\theta = \frac{2x}{D}$$

By defining a unit coefficient for $D = 1$ ft, we have

$$h_o(D) = \frac{h_o(1)}{D^{1/2}}$$

$$h_o(1) = k \left(\frac{\rho u_{\infty}}{\mu} \right)^{1/2} f(\theta)$$

The values for $\bar{h}(1)$ and $h_o(1)$ are tabulated in Table 1 for the various atmospheric models described in Table 2.

3. Experimental Results

Experiments were conducted for a flat plate and a cylinder in a low-density wind tunnel. Average heat

Table 1. Heat transfer coefficients for various Martian atmospheric models

Measuring parameter	Heat transfer coefficient, Btu/h-ft ² -°R									
	Atmospheric model									
	1	2	3	4	5	6	7	8	9	10
Flat plate: $\bar{h}(1)$	0.737	0.609	0.803	0.644	0.876	0.766	0.674	0.552	0.958	0.945
Cylinder: $h_o(1), \theta = 0$	1.17	0.966	1.274	1.022	1.392	1.217	1.07	0.876	1.520	1.50
$\theta = 20$	1.14	0.942	1.243	0.997	1.357	1.188	1.043	0.855	1.482	1.462
$\theta = 40$	1.05	0.865	1.142	0.916	1.247	1.091	0.959	0.785	1.362	1.345
$\theta = 60$	0.877	0.724	0.955	0.766	1.041	0.913	0.801	0.657	1.139	1.124
$\theta = 80$	0.603	0.498	0.658	0.527	0.718	0.628	0.552	0.452	0.785	0.775

Table 2. Characteristics of various Martian atmospheric models

Atmospheric parameter	Atmospheric model									
	1	2	3	4	5	6	7	8	9	10
Surface density, 10^5 g/cm ³	0.955	1.85	1.365	2.57	1.91	3.08	0.68	1.32	2.73	3.83
Surface temperature, °K	275	200	275	200	275	200	275	200	275	200
Specific heat of mixture, cal/g-°C	0.230	0.166	0.230	0.153	0.230	0.174	0.230	0.166	0.230	0.207
Viscosity, 10^5 g/cm-s	15.928	10.170	15.928	11.667	15.928	13.027	15.928	10.170	15.928	13.174
Conductivity, 10^5 cal/cm-s-°K	50.819	23.455	50.819	25.070	50.819	31.408	50.819	23.405	50.819	38.187
Prandtl number (at surface)	0.721	0.718	0.721	0.711	0.721	0.721	0.721	0.718	0.721	0.714
Continuous wind speed, ft/s	186.0	186.0	155.5	155.5	131.5	131.5	220.0	220.0	110.0	110.0
Maximum wind speed, ft/s	470.0	470.0	390.0	309.0	330.0	330.0	556.0	556.0	278.0	278.0
Reynolds number (based on continuous wind speed and $L = 5$ ft), $\times 10^{-5}$	0.526	1.573	0.626	1.59	0.742	1.45	0.442	1.327	0.888	1.488
Mach number (based on continuous wind speed)	0.178	0.249	0.149	0.201	0.126	0.156	0.211	0.295	0.105	0.124
Mach number (based on maximum wind speed)	0.451	0.630	0.374	0.504	0.316	0.392	0.533	0.745	0.267	0.313

transfer coefficients were measured by the following relationship:

$$\bar{h} = \frac{q_i - q_l}{A(T_s - T_\infty)} \Big|_{\text{steady state}}$$

where

q_i = main heat input to the test section

q_l = sum of the heat loss other than convection

A = test section surface area

T_s = test surface temperature

T_∞ = free stream temperature

The results were obtained from 18 test runs, the pressure ranging from 3 to 10 mm Hg and the velocity from 90 to 230 ft/s. For each test run the estimated heat losses, q_l , were tabulated as percentage of the total heat input, q_i , to the main test section. In analyzing the ex-

perimental results, lines indicating the percent deviation were included in the comparison curves (Figs. 10 and 11) to show the degree of agreement between the experimental and computed values of the mean heat transfer coefficients. Figure 10, for the flat plate, shows that most experimental values lay within a range of 10% from the computed values. The same extent of agreement could be concluded for the cylinder (Fig. 11), except that the experimental values were consistently higher than the computed values. This was probably due to the difference in the Eckert's coefficient derived for air, for which the correlations for the cylinder were deduced, where the tests were conducted for CO₂ atmosphere.

4. Conclusions

The experimental results demonstrate clearly that laminar continuum flow conditions existed over the entire range of this investigation. Hence, convective heat transfer rates in a Martian atmosphere can be obtained from existing solutions.

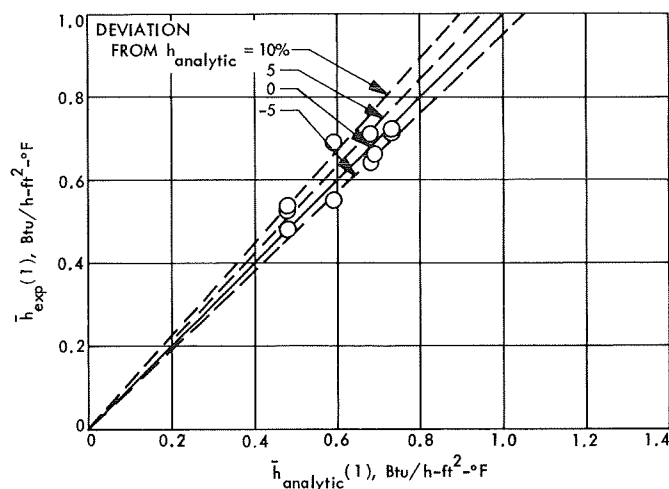


Fig. 10. Comparison of experimental and analytical mean heat transfer coefficients for flat plate

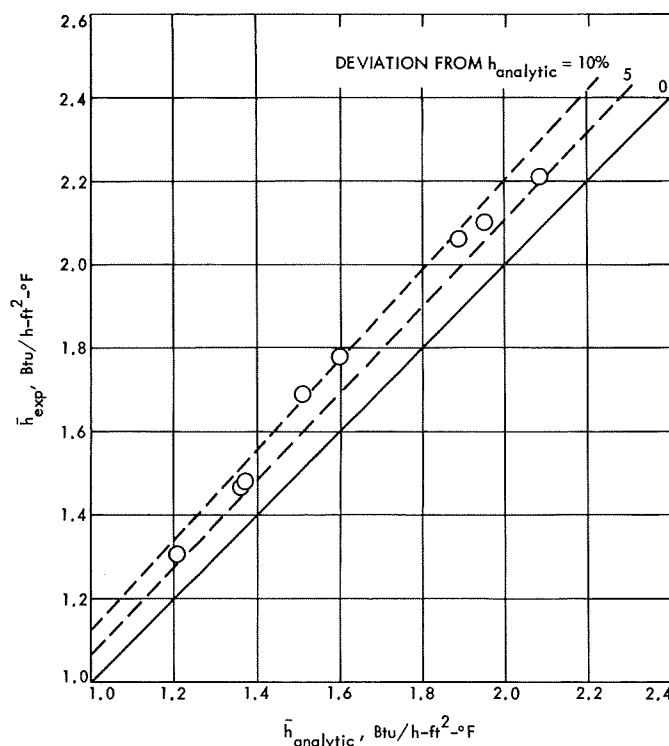


Fig. 11. Comparison of experimental and analytical mean heat transfer coefficients for cylinder

The maximum error between the measured and calculated values for models tested was of the order of 10%. The major portion of this error is attributed to the uncertainties in the estimation of the heat losses.

The difference in the Prandtl number of air and carbon dioxide (0.71 and 0.77, respectively) does not contribute

sufficiently to the results to warrant the use of a specially prepared working fluid.

References

1. Johnston, H. L., Mattox, R. W., and Powers, R. W., *Viscosities of Air and Nitrogen at Low Pressures*, NACA TN 2546. National Advisory Committee for Aeronautics, Washington, 1951.
2. Schaaf, S. A., and Chambre, P. L., *Fundamentals of Gas Dynamics*, Section H. Edited by H. W. Emmons. Princeton Series on High Speed Aerodynamics and Jet Propulsion, Vol. III, pp. 689-690, 1958.
3. Stalder, J. R., Goodwin, G., and Creager, M. O., NACA TN 2244, 1950, and NACA TN 2438, 1951. National Advisory Committee for Aeronautics, Washington.
4. Eckert, E. R. G., and Drake, R. E., *Heat and Mass Transfer*, Second Edition. McGraw-Hill Book Co., Inc., New York, 1959.

E. Development of Advanced Composite Liquid Propulsion Pressure Vessels,⁸ A. Knoell

1. Introduction

The potential weight savings and improved reliability afforded by the use of advanced composites such as boron-epoxy and graphite-epoxy in spacecraft structural applications has precipitated considerable recent interest and design development activity. Thus far, however, only a limited number of structural components have been fabricated and tested. In the area of pressure vessel technology, filament wound tanks have been fabricated and tested as reported in Ref. 1. Results obtained in that program were encouraging and essentially demonstrated that improved tank performance factors could be achieved through the use of advanced composites.

On the basis of this experience and in-house preliminary design comparison studies, a program was initiated to develop advanced composite, filament-wound, large-scale pressure vessels that would satisfy anticipated liquid propulsion system requirements on advanced unmanned spacecraft missions. These requirements are based on the use of the space storable propellants OF_2 and B_2H_6 , and a nominal systems operating temperature of -230°F . Additional motivation for considering advanced composites in this application was provided by the realization that because of the high modulus of these materials, the fuel-tank material compatibility problem can be more easily solved through the use of admissible

⁸This work was performed to support the Chemical Rocket Experimental Engineering task under the development of Spacecraft Propulsion Module Using Space-Storable Propellants for the Propulsion Division.

non-structural liners. The scope of the program consists of the design, analysis, fabrication, and testing of both pressurant and propellant prototype vessels. At this stage of development, preliminary tank design and materials evaluation studies⁹ have been completed. Program sequencing calls for the completion of final tank designs followed by fabrication and then testing of the vessels.

2. Preliminary Design Development

The design constraints for this program basically require that leak-proof, fuel-compatible pressurant and propellant tanks be designed to exhibit a burst of pressure approximately 9000 and 800 psia, respectively at a design temperature of approximately -320°F . The approach taken to satisfy these constraints was to develop tank designs using FW structural design techniques, metal-liner technology and materials characterization data. Preliminary tank designs were developed that exhibited the characteristics given below.

a. Pressurant tank design. The structural tank configuration is that of a planar-wound, 56 ply, oblate spheroid with semi-major and semi-minor diameters of approximately 13 and 8 in., respectively. The materials evaluation study indicated that the tank should be made of boron-epoxy overwrapped on an aluminum liner with aluminum boss fittings.

b. Propellant tank design. For the propellant tank, the configuration is that of a 36-in.-diam cylinder with oblate spheroid domes. The overall tank length is approximately 50 in. It was again shown that the tank material system should be boron-epoxy overwrapped on an aluminum liner with aluminum boss fittings.

The major problem area brought to light thus far centers on the development of a suitable boss-liner joint configuration. This area, as well as others relating to environmental effects on material structural response, will be explored during the process of developing the final tank design configurations.

Reference

1. Morris, E., and Alfring, R., *Cryogenic Boron Filament Wound Containment Vessels*, NASA CR 72330. National Aeronautics and Space Administration, Washington, Nov. 1967.

⁹In this program, materials support is being provided and coordinated by Mr. W. Jensen of the Materials Section.

F. Optimization of Space Antenna Concepts, E. Heer and J. N. Yang

1. Introduction

In many of its aspects, the evaluation of antenna concepts is not yet amenable to strictly quantitative procedures. Many decisions in the evaluation process are made by using good engineering judgment based on extrapolation of data available from related projects, on experience, on intuition, and the like. In this article, it is assumed that through these, or similar considerations on the systems level, certain ranges of the systems' parameters such as antenna gain, communication frequency, antenna configuration, weight, etc., have been identified as acceptable or optimum. These parameters are then used as constraints in the optimization problem, where given antenna concepts are optimized with respect to minimum cost. In the work of this article, a simplified gain equation is used that was developed by J. Ruze (Ref. 1). However, before finally deciding on using this equation, a careful evaluation of the general expressions for antenna gain, as developed in Ref. 2, has been made to assess the influence of the various assumptions on the present optimization problem. These influences are secondary and are therefore neglected.

2. General Outline

The antenna concept optimization problem considered here can be stated in the following form:

Minimize the objective function "cost":

$$C = C_1(D, N, b) + C_2(D, N) \quad (1)$$

Subject to the constraints:

$$G_1 = \eta \left(\frac{\pi D}{\lambda} \right)^2 \exp \left\{ - \left[\frac{4\pi \delta(D, N, b)}{\lambda} \right]^2 \right\} \geq G_{01} \quad (2)$$

$$G_2 = \eta \left(\frac{\pi D}{\lambda + \Delta \lambda} \right)^2 \exp \left\{ - \left[\frac{4\pi \delta(D, N, b)}{\lambda + \Delta \lambda} \right]^2 \right\} \geq G_{02} \quad (3)$$

$$\lambda \geq \lambda_0 \quad (4)$$

$$\Delta \lambda \geq \Delta \lambda_0 \quad (5)$$

$$W(D, N) \leq W_0 \quad (6)$$

where λ , $\Delta \lambda$, D , b , and N are the design variables.

This formally stated optimization problem, and the symbolism used, require some explanation. The objective function "cost" includes two parts; (1) the cost of the fabricated, delivered, and packaged antenna, and (2) the cost of lifting the antenna into outer space, and the attitude controlling it during the mission. The first cost, $C_1(D, N, b)$, is roughly a function of deployed antenna diameter D , a measure of weight per unit area N , and of a measure of fabrication tolerance b . The second cost, $C_2(D, N)$, is essentially a function of the total antenna weight $W(D, N)$, and hence, a function of D and N , where it is assumed that W is a function of D and N only and is constrained to be less than a given weight W_0 .

In many cases, operation at more than one frequency or wave length is required. The above equations reflect simultaneous operation at two different frequencies corresponding to wave lengths λ and $\lambda + \Delta\lambda$, where λ and $\Delta\lambda$ are constrained by overall system considerations to be larger than certain given values λ_0 and $\Delta\lambda_0$, respectively. The gains of the antenna G_1 and G_2 corresponding to λ and $\lambda + \Delta\lambda$ are also constrained by overall system considerations to be larger than certain given values G_{01} and G_{02} , respectively. Optimization for single-frequency operation does not require Eqs. (3) and (5), while optimization at more than two frequencies requires additional equations of the form of Eqs. (3) and (5) for each additional frequency.

The gain, Eqs. (2) and (3), is derived based on the assumption that the deviations, $\delta(D, N, b)$, from the perfect antenna shape are stationary random with gaussian distribution and that the correlated regions are small compared to the antenna diameter. Thus, $\delta(D, N, b)$ is the total rms deviation of the antenna reflector surface and is assumed to consist of essentially three contributions: (1) the rms deviation δ_m due to manufacturing errors, (2) the rms deviation δ_s due to design (e.g., a reflector mesh stretching between two ribs, thus deviating from an intended parabolic shape), (3) the rms deviation δ_c due to environmental conditions (e.g., temperature changes, loads, etc.). Some detailed considerations lead to the conclusion that these three rms deviations depend mainly on the design variables D , b , and N and that

$$\delta_m = \delta_m(D, b)$$

$$\delta_s = \delta_s(D, N)$$

$$\delta_c = \delta_c(D)$$

Hence, assuming that the three error contributions are independent, one has

$$\delta^2(D, N, b) = \delta_m^2(D, b) + \delta_s^2(D, N) + \delta_c^2(D) \quad (7)$$

The aperture efficiency η is approximately 0.6 to 0.7 and varies relatively little from antenna to antenna. This influence on the antenna optimization results is therefore neglected by assuming that η is a constant.

3. Functional Relations

The optimization problem stated by Eqs. (1)–(7) still requires the explicit specification of the functions C_1 , C_2 , δ , and W in terms of the design variables. For this purpose, data of a few existing antennas has been used and are fitted and extrapolated with sufficient free parameters so that a broad spectrum of antenna concepts is covered.

The antenna cost C_1 can then be expressed as

$$C_1 = a_1 D^{a_2} N^{a_3} b^{a_{10}} \quad (8)$$

where a_1 , a_2 , a_3 , and a_{10} are constant parameters. $a_1 D^{a_2} N^{a_3}$ is the cost of the antenna for a certain manufacturing tolerance level $b = 1$, and a_1 is a measure of the cost per unit antenna area per unit weight.

The cost C_2 can be expressed as

$$C_2 = a_3 W^{a_4} \quad (9)$$

where a_3 is the cost per unit weight of payload and W can be expressed as

$$W = N D^{a_5} \quad (10)$$

Thus, using Eqs. (8), (9), and (10), the objective function Eq. (1) becomes

$$C = a_1 D^{a_2} N^{a_3} b^{a_{10}} + a_3 N^{a_4} D^{a_4 a_5} \quad (11)$$

The rms deviation in Eq. (7) can then be represented with sufficient accuracy in the following form:

$$\delta^2(D, N, b) = (a_7 D^{a_8} / b^{a_{10}})^2 + (a_{11} D / N^{a_6})^2 + (a_{12} D)^2 \quad (12)$$

Eqs. (10), (11), and (12) contain the twelve parameters a_1 – a_{12} , which represent the physical constants of antenna concepts. These parameters can be determined from either existing antennas or past experience.

4. Conclusions

The objective function Eq. (1) and the constraint Eqs. (2), (3), and (6) are nonlinear functions of the design variables. It appears, therefore, that the optimization technique most appropriately used is a gradient move technique described in Ref. 3.

The results of the optimization process are those values of the design variables that yield minimum cost for antennas within the specified constraints. The detailed antenna design and fabrication should then be such as to satisfy these design variables.

As a next step in the present process, the above optimization will be carried out for several sets of judiciously selected parameter values based on existing antennas. A computer program will be developed for remote terminal operation and the sensitivity of antenna designs to these parameters will be determined.

References

1. Ruze, J., "Antenna Tolerance Theory—A Review," *Proc. IEEE*, Vol. 54, No. 4, pp. 633–640, Apr. 1966.
2. Silver, S., *Microwave Antenna Theory and Design*. McGraw-Hill Book Co., Inc., New York, 1949.
3. Shimozuka, M., and Yang, J. N., "Optimum Structural Design Based on Reliability and Proof-Load Test," *Ann. Assur. Sci.*, Vol. 8, pp. 375–391, July 1969.

G. A Frequency Domain Solution for the Linear Attitude Control Problem of Spacecraft With Flexible Appendages, M. R. Trubert and J. A. Garba

1. Introduction

The anticipated advent of spacecraft with large flexible appendages such as long solar panels or large antennas introduces structural modes of vibration that fall in the frequency range of the attitude-control system. Consequently, there exists a coupling between structural flexibility and the attitude-control system which imposes constraints on the study of the stability of the system. This problem has been reviewed and studied at length in Ref. 1. In this article, a non-traditional frequency domain/Fourier transform approach is proposed.

Small motion is assumed to linearize the equations. The open loop transfer function (frequency response) of the spacecraft relating attitude angle to control torques, is computed numerically in terms of discrete real frequencies over the frequency range of interest. Then the transfer function of the control loop is introduced and the transfer function of the closed-loop system relating attitude angles to a disturbance torque is computed numerically. The method consists of studying the stability of the system by the concept of the response to the unit impulse and Fourier transform. The use of real frequency ω associated with the Fourier transform technique permits a numerical computation of the time history of the response to the unit impulse by inverse Fourier transformation, giving a stability criterion by direct inspection. The Fourier transform technique is in contrast with the more traditional Laplace transform, which is usually limited to algebraic computation in terms of a complex argument S . The proposed technique has the advantage of permitting the use of a large number of elastic modes without any computational difficulties, a definite advantage over the Laplace transform method. Finally, the response to any external transient disturbance torques or force can also be computed by the same method.

2. Structural Representation

Let us consider a spacecraft composed of a rigid bus and a flexible appendage attached to the bus (Fig. 12). The appendage is attached to the bus at a number of points, which are referred to as the base of the appendage. It is natural and convenient to introduce the flexibility of the appendage in its deflection with respect to those attachment points. The corresponding modes of vibration are the "cantilever modes."

The fundamental problem of the attitude control system is to control the orientation of the spacecraft in flight with respect to an inertial frame of reference $O_1 Y_1 Y_2 Y_3$. Let us assume that it is desired that the bus has a given orientation with respect to the inertial frame. Call $O X_1 X_2 X_3$ a system of coordinates of origin O fixed in the bus (Fig. 12). We assume that the bus has only a small rotational motion with respect to the inertial frame. Call $\theta_1(t)$, $\theta_2(t)$, $\theta_3(t)$ the three time-variable rotations of the bus with respect to the inertial frame. The three translations of O are of no direct interest in attitude control, and while they are considered in the actual derivation of the equations of motion, they are not explicitly represented in the final result; therefore, they will not be shown here. Call $T_1(t)$, $T_2(t)$, $T_3(t)$ the three torques applied on the structure about the three axes $X_1 X_2 X_3$ by the control system, e.g., the control cold gas jets.

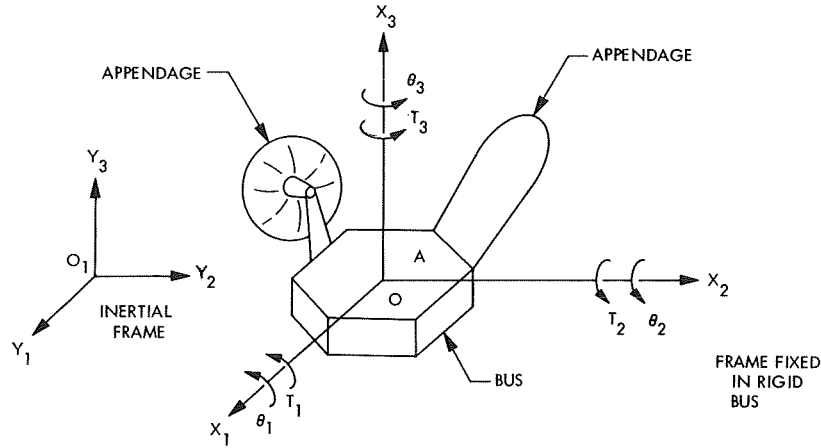


Fig. 12. Spacecraft with appendages

Now let us take the Fourier transform of each of the six variables θ 's and T 's, i.e.,

$$\bar{\theta}_j(\omega) = \bar{\theta}_j = \int_{-\infty}^{+\infty} \theta_j(t) \exp(-i\omega t) dt \quad (1)$$

$$\bar{T}_j(\omega) = \bar{T}_j = \int_{-\infty}^{+\infty} T_j(t) \exp(-i\omega t) dt \quad j = 1, 2, 3 \quad (2)$$

and form two columns

$$\{\bar{\theta}\} = \begin{Bmatrix} \bar{\theta}_1 \\ \bar{\theta}_2 \\ \bar{\theta}_3 \end{Bmatrix}, \quad \{\bar{T}\} = \begin{Bmatrix} \bar{T}_1 \\ \bar{T}_2 \\ \bar{T}_3 \end{Bmatrix} \quad (3)$$

of the rotations and the control torques.

The rotations $\{\bar{\theta}\}$ and the torques $\{\bar{T}\}$ are related by

$$\{\bar{\theta}\} = [Y]\{\bar{T}\} \quad (4)$$

in which $[Y] = [Y(\omega)]$ is the 3×3 transfer function matrix of the structure that can be obtained from the modal properties. From a system viewpoint, the structure can be represented by a black box with an input T , an output θ and a transfer function Y (Fig. 13).

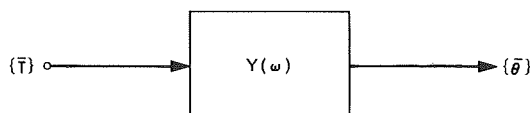


Fig. 13. Structure block diagram

3. Stability Study for Linear Control

Let us now introduce a control loop in the system by assuming that a control device senses the response of the bus, e.g., the three rotations $\{\theta\}$. The signals $\{\theta\}$ go through a control system with a transfer function matrix $[S(\omega)]$ and drive the bus with three control torques $\{\bar{T}_c\}$ (Fig. 14).

$$\{\bar{T}_c\} = \begin{Bmatrix} \bar{T}_{c1} \\ \bar{T}_{c2} \\ \bar{T}_{c3} \end{Bmatrix} \quad (5)$$

The control torques $\{\bar{T}_c\}$ are related to the angles $\{\bar{\theta}\}$ by

$$\{\bar{T}_c\} = -[S(\omega)]\{\bar{\theta}\} \quad (6)$$

In the controlling configuration, the torque $\{\bar{T}\}$ of Eq. (4) is the sum of the control torque $\{\bar{T}_c\}$ and a disturbance torque $\{\bar{T}_d\}$

$$\{\bar{T}\} = \{\bar{T}_c\} + \{\bar{T}_d\} \quad (7)$$

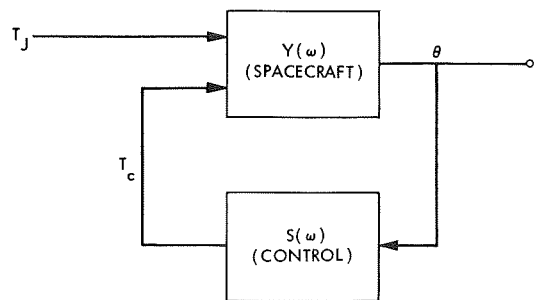


Fig. 14. Spacecraft and control block diagram

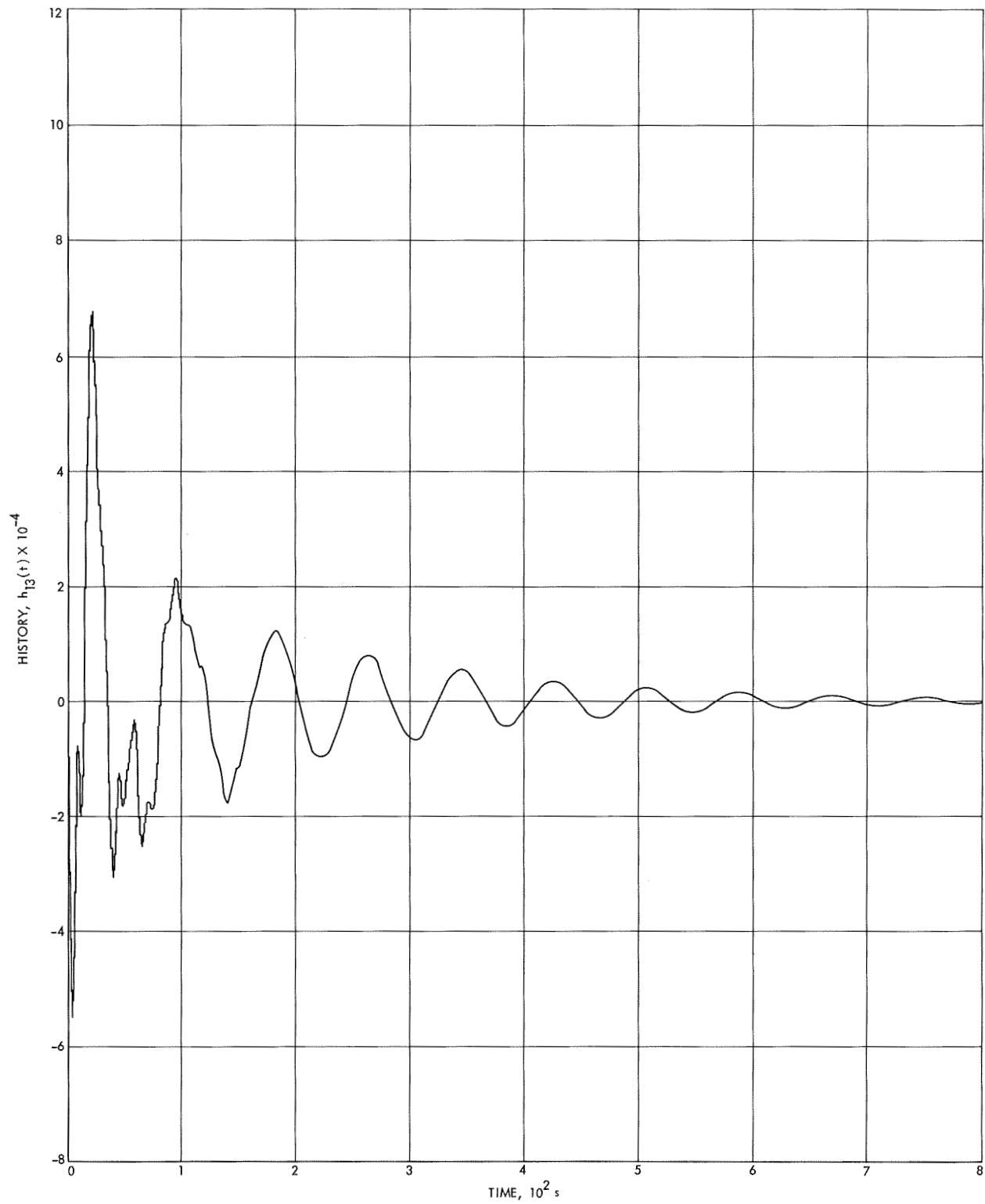


Fig. 15. Unit impulse response $h_{13}(t)$ (thrusters on bus)

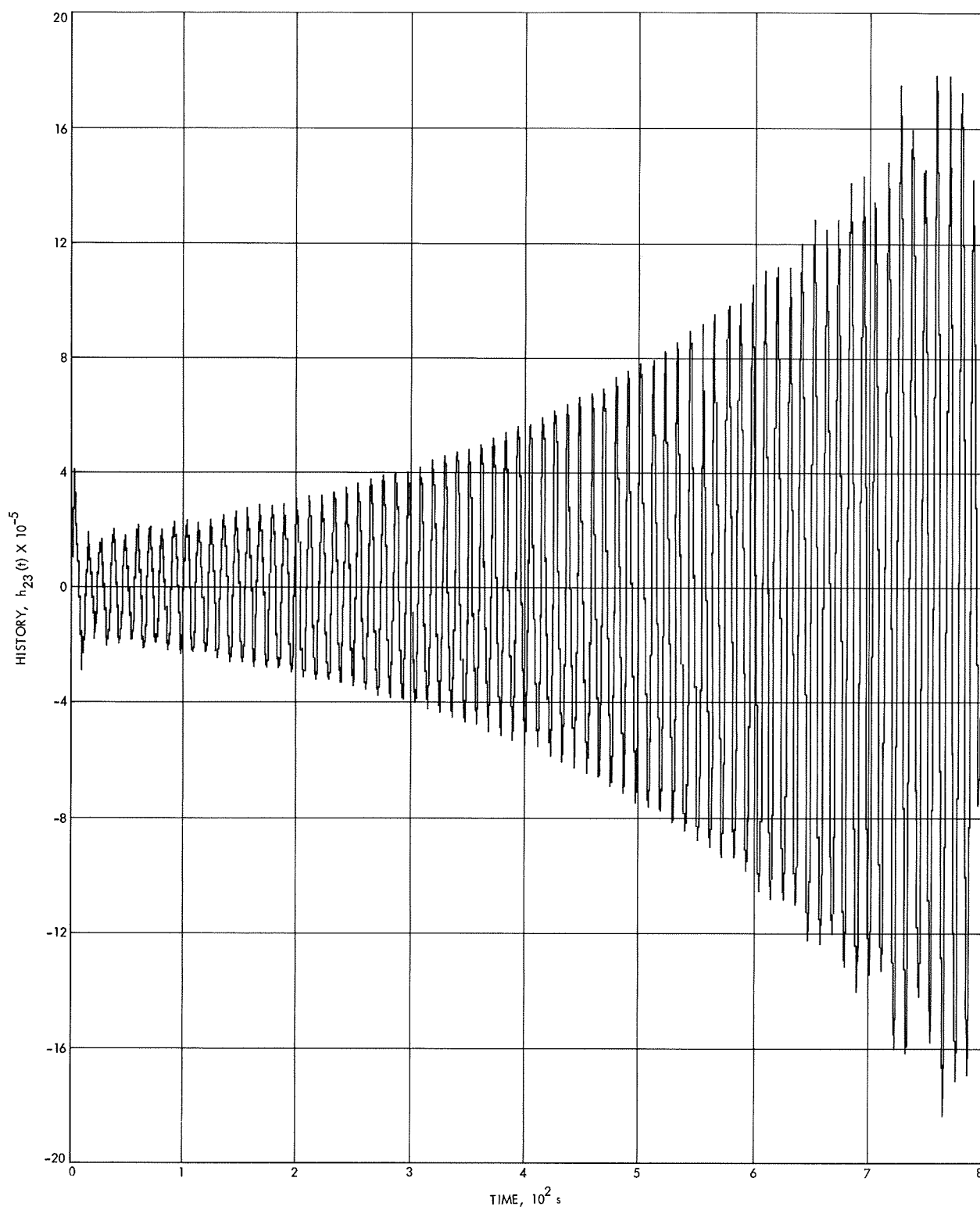


Fig. 16. Unit impulse response $h_{23}(t)$ (thrusters on appendages)

Combining Eqs. (4), (6) and (7), we obtain

$$\{\bar{\theta}\} = [1 + YS]^{-1} [Y] \{\bar{T}_d\} \quad (8)$$

which is the three-dimensional equation of the attitude control problem in the frequency domain. For short, Eq. (8) is written

$$\{\bar{\theta}\} = [\mathcal{H}(\omega)] \{\bar{T}_d\} \quad (9)$$

where

$$[\mathcal{H}(\omega)] = [1 + YS]^{-1} [Y] \quad (10)$$

is a symmetric transfer function matrix

$$[\mathcal{H}(\omega)] = \begin{bmatrix} \mathcal{H}_{11} & \mathcal{H}_{12} & \mathcal{H}_{13} \\ \mathcal{H}_{21} & \mathcal{H}_{22} & \mathcal{H}_{23} \\ \mathcal{H}_{31} & \mathcal{H}_{32} & \mathcal{H}_{33} \end{bmatrix} \quad (11)$$

$$\mathcal{H}_{\alpha\beta}(\omega) = \mathcal{H}_{\beta\alpha}(\omega) \quad \alpha, \beta = 1, 2, 3 \quad (12)$$

The matrix $[\mathcal{H}(\omega)]$ is a 3×3 frequency-dependent matrix numerically computed for a range of frequencies ω from the cantilever normal modes of the appendages, the mass properties of the bus, and the characteristics of the linear control system.

The stability of the system spacecraft/control can be studied by applying a unit impulse torque on each torquing mechanism in turn. This is equivalent to looking at the unit impulse matrix $[h(t)]$ corresponding to the matrix $[\mathcal{H}(\omega)]$. Each term of $[h(t)]$ is the inverse Fourier transform of the corresponding term of $[\mathcal{H}(\omega)]$.

$$h_{\alpha\beta}(t) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} \mathcal{H}_{\alpha\beta}(\omega) \exp(i\omega t) d\omega \quad \alpha, \beta = 1, 2, 3 \quad (13)$$

The system is unstable if any of the six different functions $h_{\alpha\beta}(t)$ is "blowing up" in time.

4. Programming

The method has been programmed on a timeshare digital computer used from a teletype terminal. It can take up to six appendages each having 10 natural cantilever modes of vibration. The orientation of the appendages with respect to the bus is a variable that can be easily modified. It has been assumed that the control torques have been obtained from cold gas jets and can be placed either on the bus or on the appendages. The

control circuits are represented by second-order polynomials in ω . The functions $h_{\alpha\beta}(t)$ are plotted in function of time on the terminal plotter. Variation of control gain can easily be made.

5. Example

The method has been applied to an idealized spacecraft consisting of a bus with two appendages each having four degrees of freedom, i.e., there is a total of eight cantilever modes. Two cases were investigated as follows:

(1) **Case 1.** The control jets were placed on the bus. Figure 15 shows a plot of one term, $h_{13}(t)$, of the matrix $[h(t)]$ indicating that for this case, the system was stable. The five other plots $h_{11}(t)$, $h_{12}(t)$, $h_{22}(t)$, $h_{23}(t)$, and $h_{33}(t)$ are not described here but exhibit a similar behavior.

(2) **Case 2.** The control jets were moved from the bus to the appendages. Figure 16 shows a plot of one of terms, $h_{23}(t)$, of the matrix $[h(t)]$. The plot indicates instability since the amplitude "blows up" in time. Similar instability was present in the five other plots of $h_{\alpha\beta}(t)$.

6. Conclusion

The foregoing two cases demonstrate that the frequency domain approach is valid to determine stability of the attitude control of spacecraft with large flexible appendages. Preliminary running time is about 20 min CPU time to calculate and plot a set of six curves for $h_{\alpha\beta}(t)$. Work is in progress to improve the speed and the capability of the method.

Reference

1. Likins, P. W., *Dynamics and Control of Flexible Space Vehicles*, Technical Report 32-1329. Jet Propulsion Laboratory, Pasadena, Calif., Feb. 15, 1969.

H. Cabling Heat Loss at Neptune, L. D. Stimpson

1. Introduction

The Thermoelectric Outer Planet Spacecraft (TOPS) will encounter extremely cold conditions at Neptune. The current spacecraft design concepts include a dozen or so external scientific instruments that must be maintained at internal temperatures near 0°F. The spaceframe being exposed to cold space will reach temperatures in the range of -350°F to -370°F. The large temperature differential between compartment interior and exterior across the attachments, coupled with the radiative losses of the instrument compartment to cold space, presents

a difficult thermal control requirement. Electrical power will be furnished by Radioisotope Thermoelectric Generators (RTGs), where the output diminishes slowly with time. Demands will be so great from the other subsystems, that there may be insufficient power to heat the instruments electrically near the end of the mission. The electrical requirement can be reduced by several means, i.e., using other power sources or heat distribution techniques; however, better isolation contributes directly to this power reduction.

Heat will leak through insulation surrounding an instrument, through support structure, and through electrical cabling. Attention is being given to the first two heat leak sources; the subject of this article is the loss through the cabling. Heat is radiated to space from the cable and is conducted to the cold spaceframe through attachments.

2. Cable Description

Wire commonly used in JPL spacecraft cabling consists of three sizes: 24 American Wire Gauge (AWG) for signal wires, 22 AWG for power leads, and 20 AWG for ground shielding. Each of these wires consists of 19 strands of 36, 34, and 32 AWG, respectively. The insulation around each wire, or bundle of 19 strands, is a fluorocarbon (TFE) with a nominal thickness of 9 mils.

Different size cables are made up according to the requirements of each instrument. Three cables have been selected that represent the spread from smaller instruments, such as the magnetometers, to larger instruments, such as the TV cameras. Characteristics for these representative cables are given in Table 3.

Table 3. Cable characteristics

Cable size	Instrument size		
	Small	Nominal	Large
Signal wires	10	18	30
Power wires	1	2	2
Ground wires	4	4	8
Cable diameter, in.	0.25	0.30	0.40
Area/length, ft ²	0.0654	0.0786	0.1048
Conductivity, Btu-ft/h-°F	0.012	0.017	0.030

The cables are wrapped with aluminized mylar resulting in an effective emittance of 0.12 using the internal temperature of the wire.

3. Analytical Results

A discrete 8-node mathematical model was used to represent cables attached to cold structure at different lengths from a warm instrument. Figure 17 shows the resulting temperature distributions along the nominal cable (Table 3) attached at the 1-ft station, 2-ft station, or 4-ft station. The heat loss from the instrument is directly proportional to the slope of these curves at $x = 0$, where the maximum shown is 1.56 W. The predominant heat transfer mode is conduction rather than radiation to space, as is evident from the near linearity of the curves. Moving the attachment further from the instrument significantly reduces the heat loss, but presents some dynamic problems. Reducing the heat conduction path through the attachment, indicated by the dashed curves in Fig. 17, is a more promising approach.

A phenomenon of higher conductivity results in the copper wire as the temperature reaches the cryogenic range. This effect is shown in Fig. 18. This effect peaks near -400°F and declines below that temperature. The high conductivity effect ratio K_{HI}/K_0 , where K_{HI} refers to high (cryogenic) conductivity and K_0 refers to normal

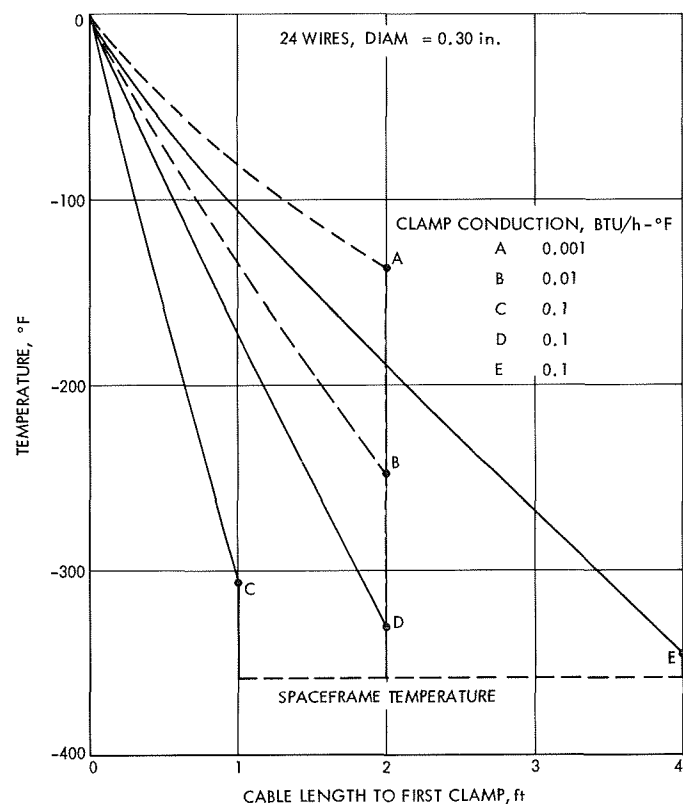


Fig. 17. Temperature distribution for nominal cable

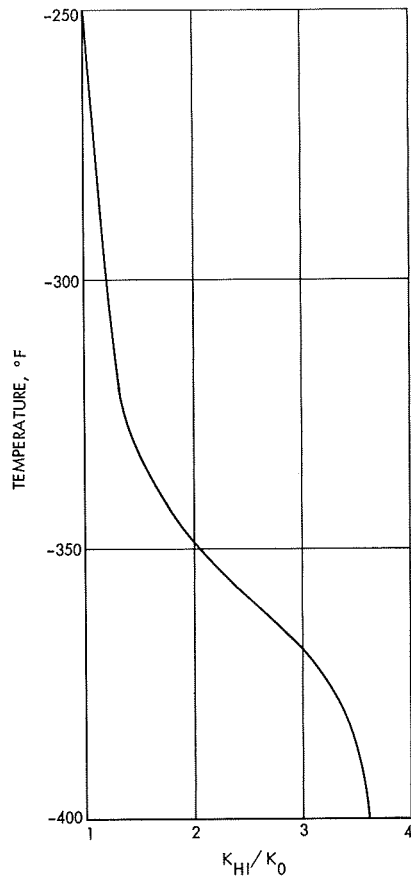


Fig. 18. High conductivity effect ratio

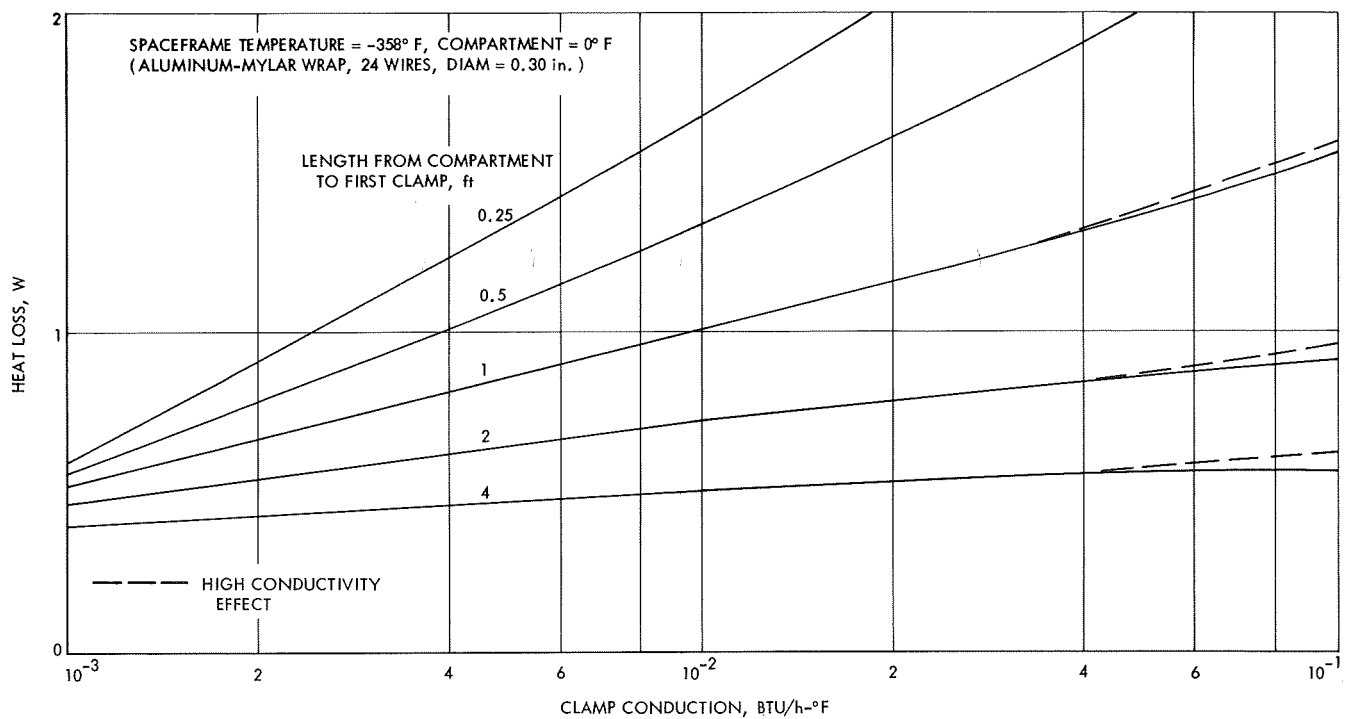


Fig. 19. Heat loss through nominal cable at Neptune

conductivity, is plotted and shows that the cable and temperature can drop into the cryogenic range, thus increasing the heat loss.

The heat loss through the nominal cable is shown in Fig. 19. It is seen that losses can be kept down below 1 W by either increasing the distance to the first clamp and/or by reducing the clamp conduction. Even a close attachment within inches can be considered for high isolation clamps. Note the small influence of the higher copper wire conductivity (dashed curves). This effect can become more important as the spaceframe becomes colder.

The heat losses for the small and large cables in Table 3 are similar to the losses shown for the nominal cable in Fig. 19, being reduced by 20% and increased by 44%, respectively. The larger cable heat losses require better isolation clamps and/or longer attachment distances.

4. Thermal Isolators

Standard cable attachments used on previous JPL spacecraft have fairly large conduction levels. Figure 20 shows a design used on *Surveyor* that has a conduction value of about 0.05 Btu/h-°F. A proposed design for a thin structural attachment that could reduce this by a factor of 10 or more is also shown. This design depends upon the cable carrying most of the load along its length and is designed for structural buckling loads. High strength composite materials such as a boron fiber or a glass polyimide offer considerable promise.

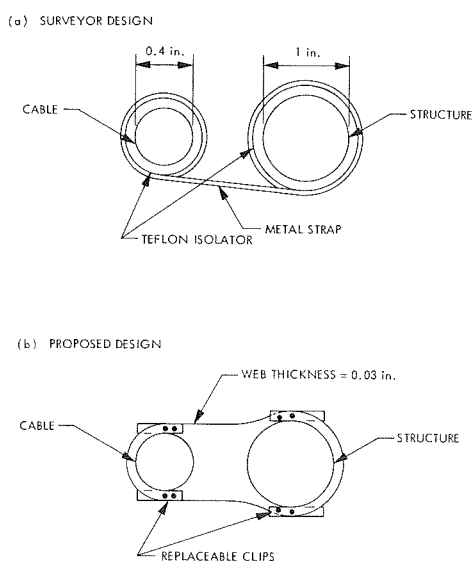


Fig. 20. Cable attachment designs

I. Gamma Ray Spectrometer Preliminary Thermal Tests, S. W. Petrick

1. Introduction

A gamma ray spectrometer (GRS) is being built at JPL for the Manned Space Center for use on future *Apollo* flights to measure gamma ray spectra and fluxes in the vicinity of the moon. Temperature control design of this instrument will provide a safe thermal environment for its critical components and provide the optimum operating temperature range during the mission.

This article summarizes the test results for two preliminary thermal tests on a prototype detector assembly (the gamma-ray sensitive portion of the instrument). The tests were started on July 24 and August 15, 1969. These starting dates are used to identify the tests.

The results of these tests were used to:

- (1) Refine the analytical model.
- (2) Gain insight into the thermal characteristics of the hardware, i.e., thermal capacitances and temperature gradients.
- (3) Choose first cut heater size.
- (4) Provide background for more elaborate future tests.
- (5) Test out efficiency of the temperature control system.

2. July 24 Test

The objectives of this test were to:

- (1) Determine the heater size required for the GRS detector.
- (2) Find the operating characteristic and operating limits of the temperature control circuit (TCC).
- (3) Determine some thermal characteristics of the detector assembly.

a. Test configuration. The tested configuration is as shown in Fig. 21. The test unit was suspended inside a vacuum chamber by nylon cords to minimize thermal conduction of heat from external sources. The wall temperature of the vacuum chamber for this test was maintained at -320°F . Warming of the unit was accomplished by the attached heater while all cooling was accomplished by radiation to the cooled chamber walls. A TCC controlled the heater attached to the detector housing.

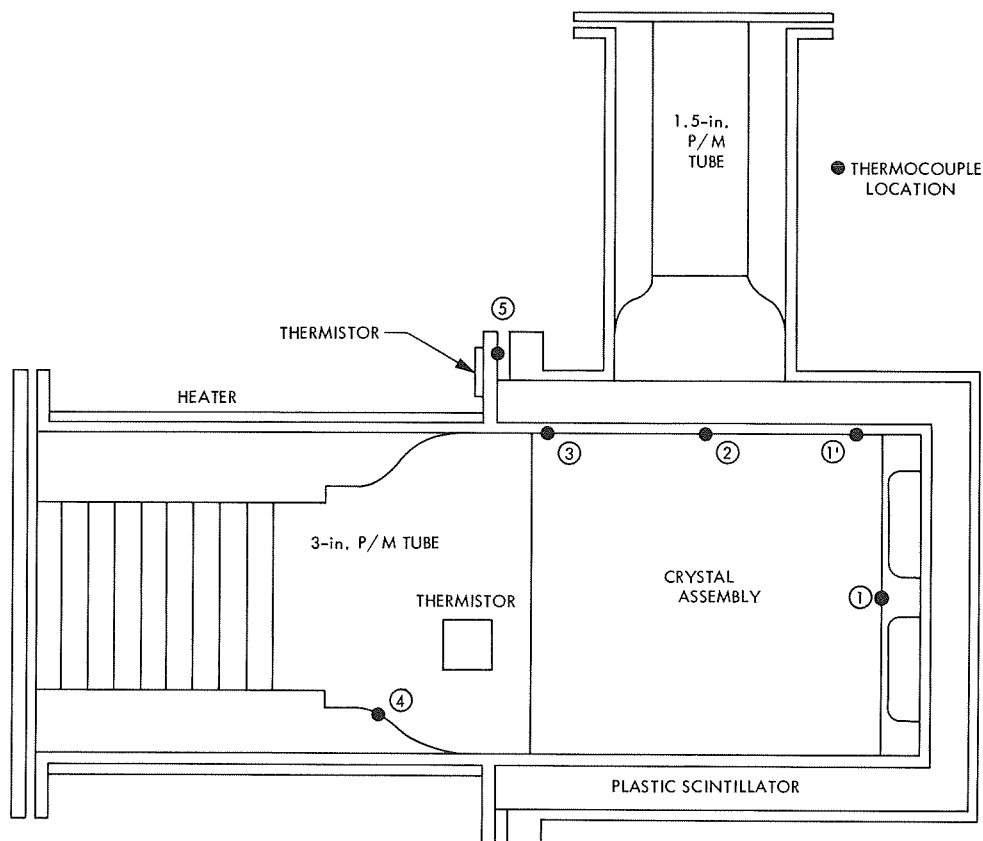


Fig. 21. Test configuration and thermocouple location

b. Instrumentation. Chromel Alumel thermocouples were located as indicated in Fig. 21. The two thermistors were used to provide input to the temperature control circuits. Thermocouple millivolt values were displayed on a Honeywell millivolt recorder.

c. Test procedures. The unit was first cycled with three different heater power inputs (Figs. 22 and 23). For the first cycle, 10 W of electrical power were input to the heater until the temperature of the unit rose to 125°F. The unit was allowed to cool to a temperature of 32°F by radiating to the liquid-nitrogen-cooled walls of the vacuum chamber. For the second cycle, 15 W were applied to the heater and 20 W were applied to the heater during the third cycle. After the temperature cycling, the heater control circuit was activated. The temperature of the unit was then allowed to approach that of the TCC set point.

d. Results. Lack of time and funding imposed some serious restrictions on the number and locations of the thermocouples used. Direct temperature measurements were not taken on critical thermal control components. No thermocouples were located in or on the NaI crystal,

although thermocouples were located on the can containing the crystal. Unfortunately, there is no information available to correlate the temperature of the NaI crystal with that of the can. With the arrangement of thermocouples used (Fig. 21), there was no way that gradients in the plastic scintillator or photomultiplier (P/M) tubes could be measured.

Upon disassembly of the test unit after this test, it was discovered that the plastic scintillator was cracked. It is not known at what time during the test the scintillator cracked, but it is suspected that the high heating rates with the 20-W heater input may have been responsible (Table 4). Based on the cracked scintillator and the fact that the heating rate of a 10-W power input did not exceed the crystal manufacturer's maximum recommended rate of 10°C/h, a 10-W power level is initially suggested for the heater.

A warm-up time of 4.1 h from the lower limit of -4°F to the lower desired operating limit of 41°F is the trade-off price for the safe warm-up rate.

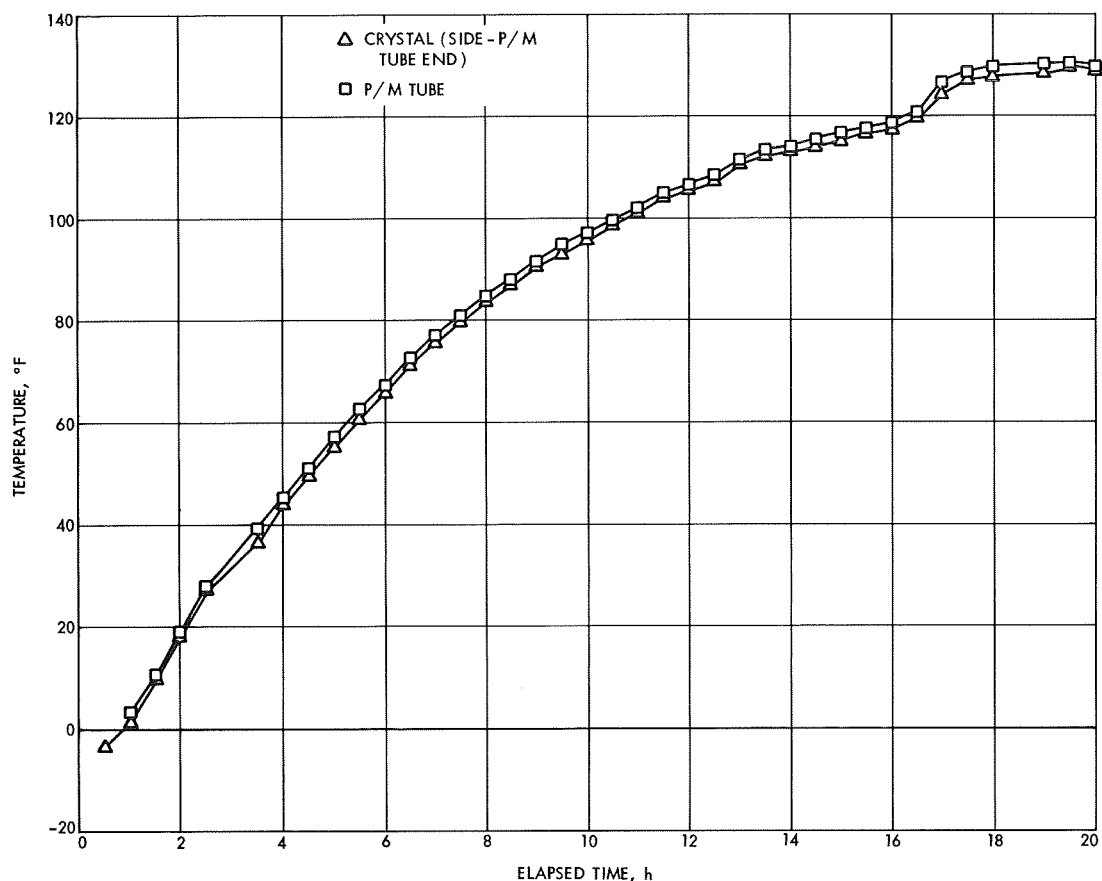


Fig. 22. Temperature response to 10-W heater power input

Table 4. Detector warm-up rates at 60°F

Heater power, W	Warm-up rate		Time from -4 to +41°F, h
	°C/h	°F/h	
10	6.1	11.0	4.1
15	12.6	22.7	2.0
20	19.2	34.6	1.3

Using 20-W heater power, the temperature control unit maintained an average temperature (measured by the thermocouples attached to the crystal assembly) of 65°F with a controlling band of $\pm 2^\circ\text{F}$. Heater-on times ranged from 1 to 6 min with a typical on-time of 3 min. The heater duty cycle was about 33%.

Based on the above data, it was decided that the heater on and off times were too erratic and that the controlling range could be increased (the desired operating temperature range of the GRS is between 41°F and 77°F).

Therefore, the temperature control unit was taken back to the vendor for alterations.

3. August 15 Test

The objectives of this test were to:

- (1) Retest the temperature control circuit.
- (2) Determine additional thermal characteristics of the detector assembly.

These objectives were accomplished by exposing the GRS to the following three sets of thermal environments:

- (1) Warm up a cooled unit until the TCC acquires control and allow to stabilize.
- (2) Cycle the walls of the vacuum chamber to simulate the variations of the thermal shield in lunar orbit.
- (3) Cool a warmed unit until the TCC acquires control and allow temperatures to stabilize.

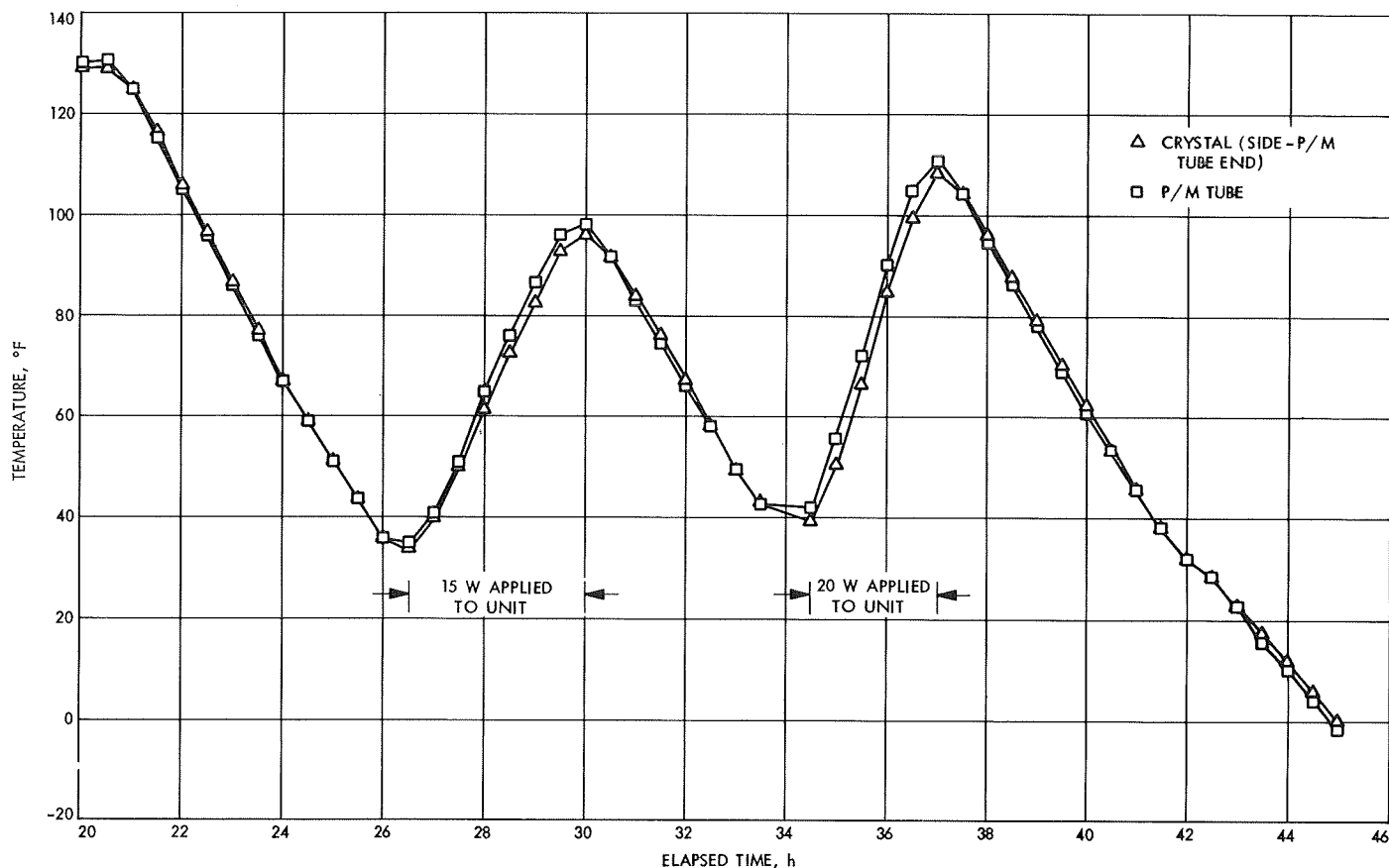


Fig. 23. Temperature response to 15- and 20-W heater power inputs

a. Test configuration and instrumentation. The configuration of this test was identical to that of the previous test with the exception of the location of thermocouple 1. This thermocouple was relocated to position 1' (Fig. 21). Unfortunately, all data on thermocouple 1' was lost because of shorted lead wires.

b. Test procedure. The first part of this test was the warmup. The walls of the vacuum chamber were maintained at -320°F and the unit was allowed to cool to 0°F . At this point, the TCC was activated. The TCC supplied the heater with a constant power of 10 W until the set temperature was achieved, at which time the heater was cycled by the TCC to maintain the unit at an average temperature of 60°F . The duty cycle of the heater during this time was 63% (Fig. 24).

After 8 h of operation at the TCC control temperatures, cycling of the vacuum chamber walls commenced. With no changes from the previous portion of the test, the walls were caused to cycle every 2 h between -60 and $+60^{\circ}\text{F}$ (Fig. 25).

After three cycles, the unit was then warmed to approximately 135°F . The TCC was then reactivated. The unit cooled to the TCC set temperature where it remained for the final 2 h of the test (Fig. 26).

c. Results. For these tests, the variation of the temperatures of the thermocouples attached to the crystal assembly was $\pm 4^{\circ}\text{F}$ about the average. Heater on-times ranged from 20 to 40 min with a typical on-time of 30 min. From this data, the TCC stability problem encountered in the test of July 24 appears greatly improved. It was noted during this test that the average controlling temperature of the TCC varied during each of the three sets of thermal environments. For set 1, the average temperature was 70°F ; for set 2, the average temperature was 60°F ; and for set 3, the average temperature was 62°F (Figs. 24, 25, and 26). Although the controlling temperature appeared to vary with the environment, it was not consistent. Since the environments of sets 1 and 3 were equivalent, the controlling temperature should have been equivalent if the controlling temperature did indeed vary directly with environment. It is

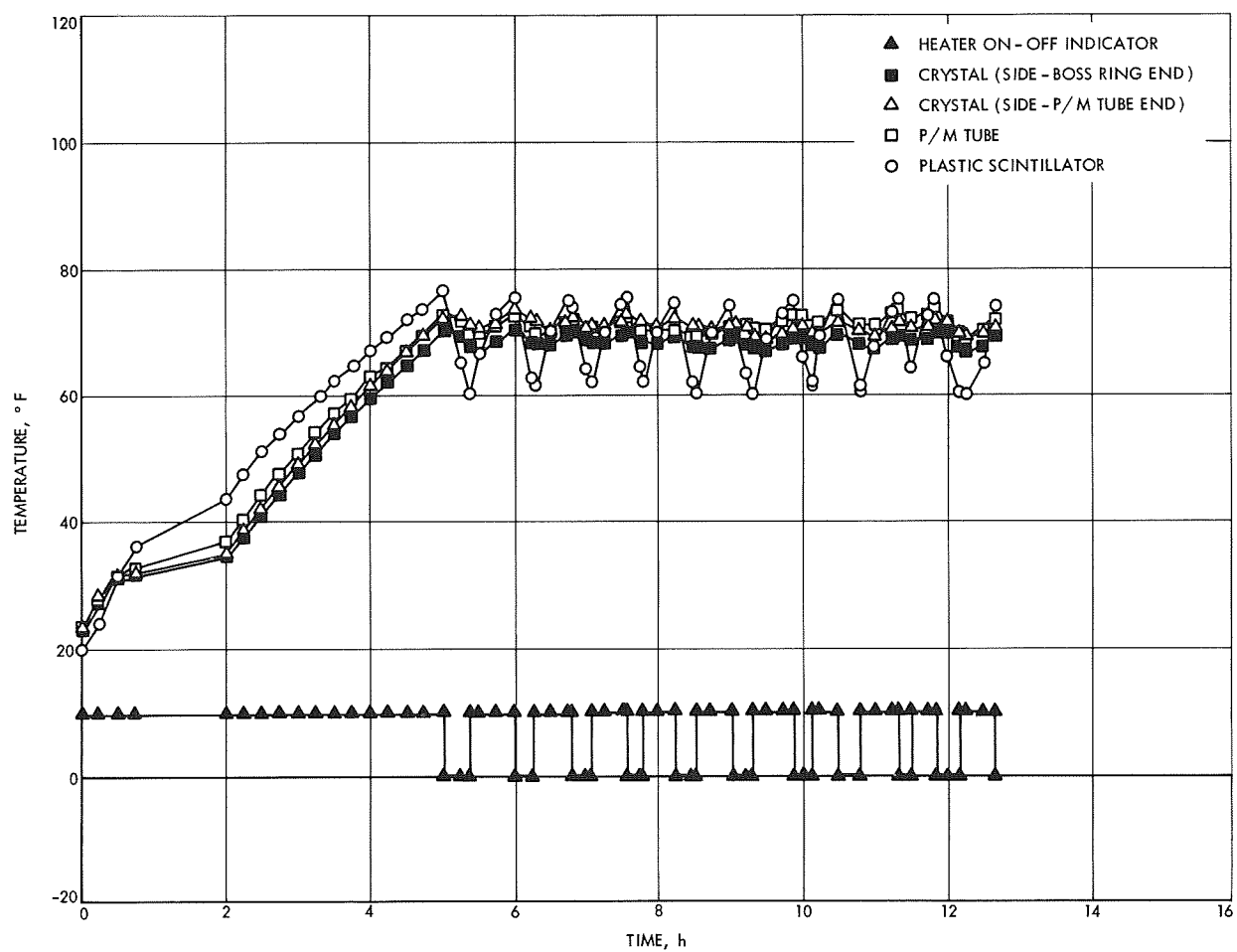


Fig. 24. Warm-up test temperature history

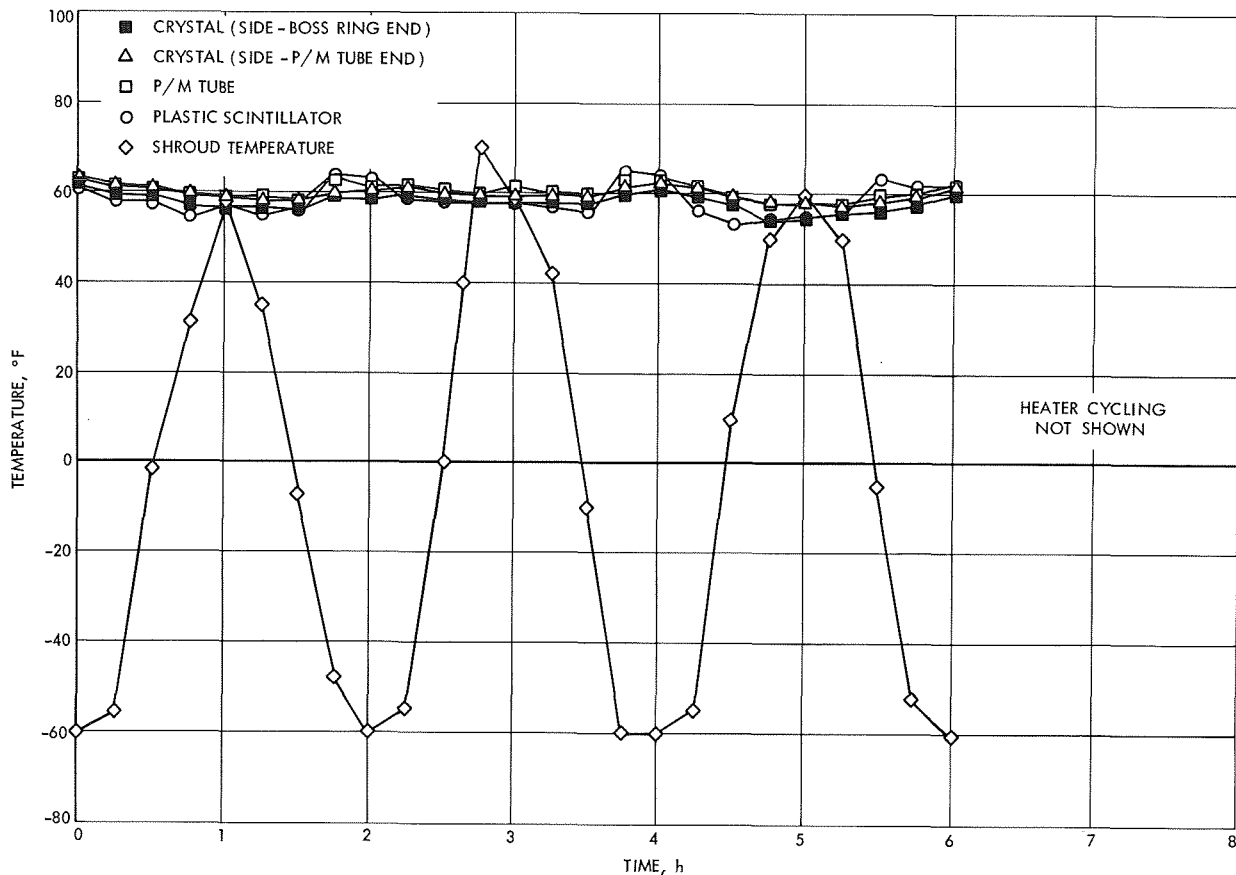


Fig. 25. Cycling test temperature history

suspected that the two-sensor controlling systems presently utilized by the TCC may be the cause of the shift. This effect will be examined in future thermal tests.

Even with the shift of the TCC temperature set point, the temperatures of the thermocouples attached to the crystal assembly remained within the 41 to 77°F range (Fig. 27).

Including the above temperature shift, the temperatures of the detector assembly appear to be very stable during the temperature cycling portion of the test. Close examination of Fig. 25 shows that the temperature variations of the detector assembly were out of phase with the temperature variations of the chamber walls. This out-of-phase condition is advantageous since it tends to minimize thermal oscillations of the detector assembly. Whether this out-of-phase condition is the result of the previous shifting of the set point with environment or a thermal character of the unit has not been determined. This condition will be covered by future tests.

The thermal capacitance of the test unit was calculated (Table 5) and compared to that used in the analytic model. The average test value is 25% higher than the analytical model.

Table 5. Thermal capacities calculated from cool-down curves

Reference curve	Radiation coefficient, $\text{Btu/h}\cdot^{\circ}\text{R} \times 10^{-10}$	Thermal capacitance, $\text{Btu}/^{\circ}\text{F}$
Fig. 21 first cool-down	2.9234	1.102
Fig. 21 second cool-down	2.9234	1.422
Fig. 21 third cool-down	2.9234	1.392
Fig. 25 cool-down	2.76	1.143
Average value of thermal capacitance = 1.265 A thermal capacitance value of 0.9932 was used in the analytical model.		

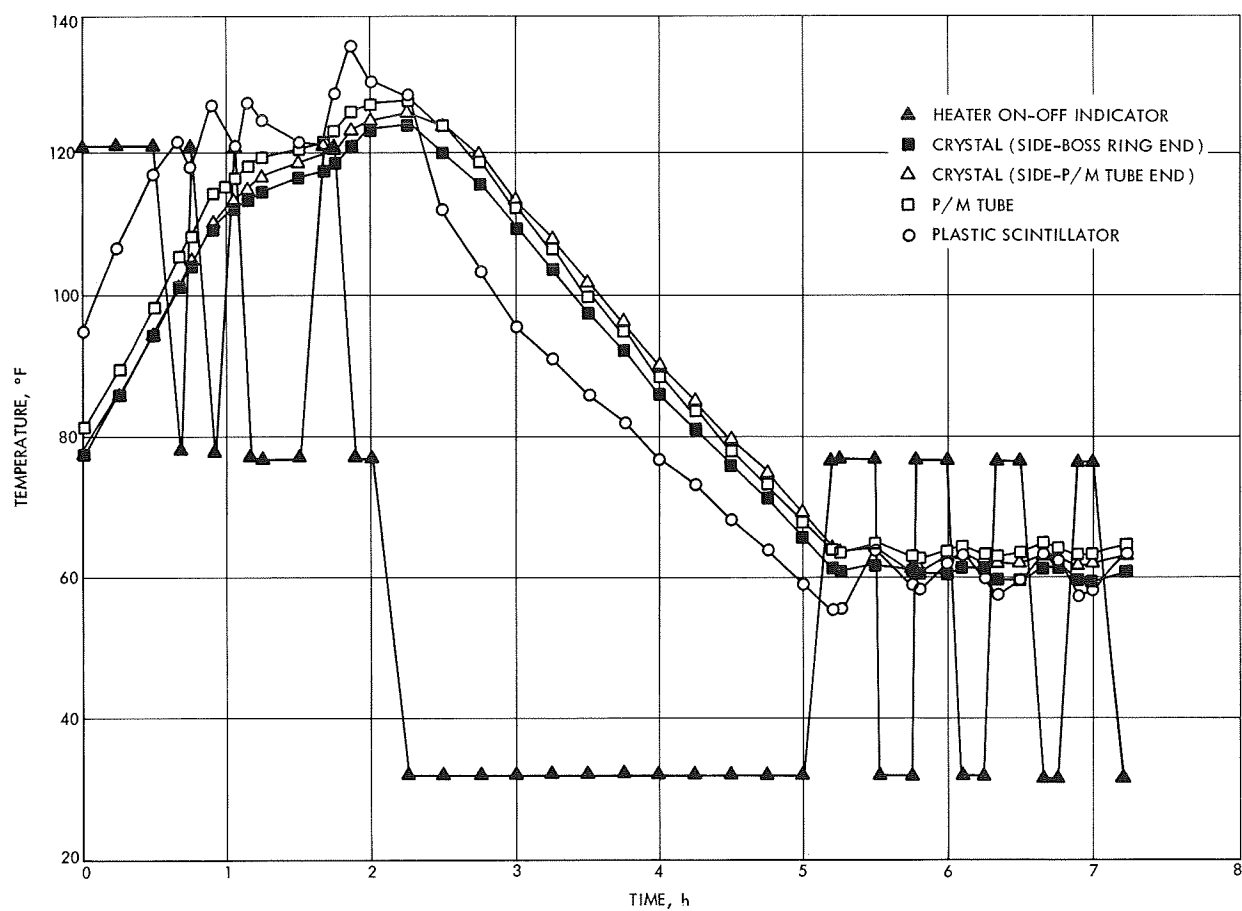


Fig. 26. Cool-down test temperature history

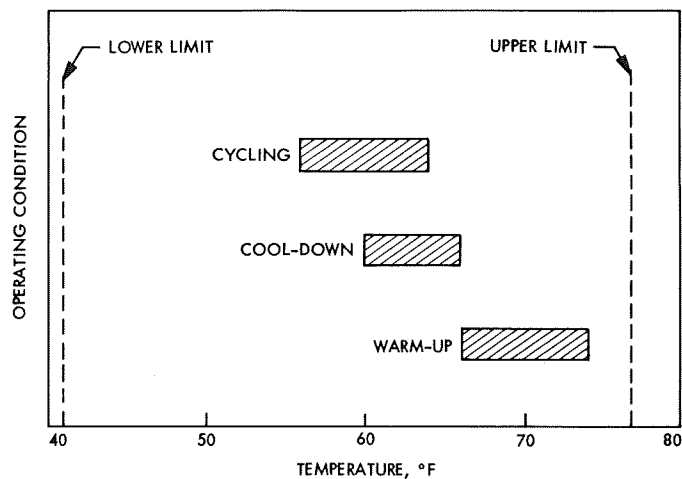


Fig. 27. Comparison of steady-state temperature control ranges for warm-up, cycling, and cool-down tests

The average cool-down rate of the four cool-down curves is 18.87°F/h. This corresponds to a cool-down time from an upper limit (120°F) to the upper operating limit (77°F) of 2.3 h.

Table 6 presents the temperature differences between the thermocouples for different operating conditions.

The thermocouples reacted as expected. The smallest temperature differences were produced while the TCC was controlling during the steady-state portions and the temperature cycling portion of the test. The temperature differences increased with increased heater power. The temperature differences for the 10-W heater power phase and the cool-down are similar.

Table 6. Temperature differences between thermocouples located on detector components

Operating condition ^a	Thermocouple location			
	Crystal assembly case to housing	P/M tube to housing	Along crystal assembly case	P/M tube to crystal assembly case
10-W power warm-up	15	12	2	3
15-W power warm-up	28	21	4	7
20-W power warm-up	39	29	5	10
Typical cool-down	17	14	2	6
Temperature cycling (heaters controlling)	6	4	1	3
Steady state (heaters controlling)	12	8	3	5
^a All temperatures are in °F.				

XV. Electronics Parts Engineering

ENGINEERING MECHANICS DIVISION

A. Matrix Test of Sterilizable Piece Parts,¹ K. Martin

1. Introduction

The object of this test conducted by Litton Systems, Contract 951706, is to support the NASA thermal-sterilization policy by studying the effects of the NASA-recommended temperature-time combinations to effect maximum electronic parts reliability. Groups P, M, I, E, and C are included among the recommended temperature-time heat-sterilization environments. The other higher stress groups were included for additional information. The temperature cycling is followed by a 10,000-h life test. The two capacitor types tested were: (1) Type 1—Sprague 350D, 39 μ F, 35-Vdc, solid tantalum, and (2) Type 2—Aerovox V423XP, 1 μ F, 200-Vdc Mylar. These types were selected from past test information, which indicated they would yield the maximum information. It would have been desirable to test more types, but funds prohibited this increase. An outline of the test program is as follows:

(1) Pretest screening.

- (a) Assign serial numbers 1-720 for each type.
- (b) Measure the parameters, screen the parts for 168 h with power applied, and remeasure the parameters.

(2) Start testing.

- (a) For each type, randomly select 540 parts that passed step (1b), above.
- (b) For each type, randomly select 30 parts for each group (matrix cell).
- (c) Measure the parameters and record the data.
- (d) Test matrix (Table 1).
- (e) Subject each group to the above specified conditions for one cycle with no power applied.
- (f) Measure the parameters and statistically analyze the data. Repeat the temperature exposure and measure the parts, both five times.
- (g) Life test the parts at 85°C with power applied. Remove the parts, measure the parameters, and statistically analyze the data at 100, 250, 500, 1000, and 2000 h.
- (h) Prepare 2000-h interim report.
- (i) Continue the life test and measure the parameters at 4000, 6000, 8000, and 10,000 h.
- (j) Prepare final report.

At the applicable measurement steps, the data were analyzed using linear regression, analysis of variance, and other statistical methods.

¹This is the final 10,000-h report which updates SPS 37-55, Vol. III, pp. 154-156.

2. Test Results

a. Results at completion of the test program matrix phase. Linear regression was used to test the significance of the change (trend) of the parameters for each group during the six cycles. Groups C, E, I, M, and P are the temperature-time stresses that JPL considers equally effective for heat sterilization. The other groups have been included in the test program so that higher temperature-shorter time stresses could be evaluated for future reference.

Type 1. There were no catastrophic failures. The capacitance did not change significantly. Eight parts exceeded the 3% dissipation factor (DF) limit. The DF failures by groups were:

Group	Failures
C	1
D	3
G	2
H	1
R	1

Groups D and G exhibited an increasing DF trend during the six cycles. The 10 μ A leakage-current limit was exceeded by 144 parts. The failures by groups were:

Group	Failures
B	1
D	12
F	5
G	18
H	12
I	1
J	12
K	20
L	16
N	11
O	14
Q	2
R	21

All groups, except Group P, exhibited a significantly increasing trend in leakage current during the six cycles.

Type 2. There was one random catastrophic failure in Group G where a lead separated from the foil element. There was one capacitance failure in Group Q. All groups except Group Q exhibited a decreasing capacitance trend during the six cycles. However, the parts remained well within the 10% tolerance limit. There were no DF failures. Groups B, C, D, E, F, G, H, I, M, N, and P exhibited a significantly decreasing trend in DF during the six cycles. There were no insulation resistance (IR) failures. Groups J, L, and R exhibited a significantly decreasing trend in IR during the six cycles, although they remained well within tolerance.

b. Test results at completion of 10,000 h of life testing.

Type 1. There were no catastrophic failures during the 10,000 h of life testing. None of the parts exceeded the capacitance tolerance limits. However, 19 of the parts exceeded the DF limits. The DF failures listed by groups are:

Group	Failures
B	1
C	1
D	6
F	1
G	4
H	1
J	1
K	1
R	3

All but 16 of the 144 matrix-testing dc leakage failures recovered as of 10,000 h of life testing. The dc leakage failures listed by groups are:

Group	Failures
B	1
D	2
H	1
L	1
M	1
N	1
P	1
Q	8

Table 1. Test matrix

25°C	105°C	115°C	125°C	135°C	145°C	160°C	Time, h
Group A 30 parts (typical) ^a						Group B	3
				Group E	Group C	Group D	9
					Group F	Group G	22
					Group H		36
			Group I	Group J	Group K		53
				Group L			92
		Group M	Group N	Group O			132
	Group P	Group Q	Group R				336

^aGroup A is the control group and does not receive heat cycling.

Type 2. There was one random catastrophic failure during the 10,000 h of life testing. There were no capacitance failures. One part in Group F exceeded the DF limit. No IR failures were noted.

3. Conclusions

a. Type 1. The capacitance of all groups has remained stable throughout the matrix and life tests; capacitance does not appear to be a sterilization testing problem. As was expected from past testing experience, the dc leakage failures tended to recover during life testing. In addition, 8 failures occurred in Group Q, which is a low-stress group. There has been no satisfactory answer for these failures. If funding permits, this stress test should be repeated with new parts. The other leakage current failures were randomly scattered throughout the matrix with no statistical significance. There were 19 DF failures. Eleven of the DF failures occurred in the 160°C groups with the other failures randomly scattered throughout the matrix. This is statistically significant and indicates that 160°C should not be used for sterilization of this type of capacitor due to DF degradation.

b. Type 2. There have been no significant catastrophic or parametric failures as of 10,000 h of life testing. There is no measurement evidence that this type of Mylar capacitor is significantly degraded from the effects of sterilization testing, although subsequent physical examination revealed that the Mylar dielectric was embrittled by the 160°C exposure.

Based on the study effort, the following table ranks the recommended sterilization stress levels in an increasing order of degradation:

Group	Temperature, °C	Time, h
P	105	336
M	115	132
E	135	22
I	125	53
C	145	9

XVI. Advanced Projects Development

ENGINEERING MECHANICS DIVISION

A. TOPS Boom Deployment Mechanisms,

E. E. Sabelman

1. Introduction

The Thermoelectric Outer Planet Spacecraft (TOPS) possesses two sets of equipment—the radioisotope thermoelectric generators (RTGs) and the science platform—which must be situated as far apart as possible. To keep the spacecraft diameter within the launch vehicle shroud envelope, the RTGs and science instruments are mounted on hinged rigid booms which are deployed after the shroud is jettisoned. The purpose of the present article is to summarize the progress in analysis and development of mechanisms for deploying the RTG and science booms.

While the problem is basically similar to the deployment of solar panels on the *Mariner* spacecraft, the moment of inertia of the RTG boom is more than two orders of magnitude greater than that of the *Mariner* Mars 1969 solar panels. Presently planned RTGs have a weight of 225 lb and a center of gravity nearly 6 ft from the hinge axis. Since the weight allocated to boom deployment mechanisms is only slightly larger than for solar panel mechanisms, it was decided to re-examine from first principles the problem of unfolding, slowing, and locking a swinging boom.

2. Analog-Physical Design Iteration

The general question of “deployment” was approached along two routes that converge on a solution or solutions to the problem. One approach was to generate a variety of applicable concepts, and select those meeting criteria of minimum weight, maximum force/weight ratio, minimum non-usable stored energy, and maximum reliability. Other criteria specific to the TOPS application were the possibility of incorporating the actuator/damper into the existing structure and the requirement for testing the boom mechanisms in the earth’s gravity field.

The second approach was to develop an analog model for a generalized swinging boom. Terms were added to the standard spring-mass-damper equation for bearing friction torque (T_B) with friction coefficient (f_B) a function of velocity ($\dot{\theta}$), and for the initial jerk or push (T_{IN}) as the boom is released by a pyrotechnic pin puller. The equation was further modified to include ground test effects from hinge axis misalignment ($T_{\phi_{\perp}}$) and from cantilever load on the bearings. A number of viscous dampers are nonlinear with velocity (C_D proportional to $\dot{\theta}^{-1/2}$) and provision for this effect was included in the model.

Typical deployment actuator torques, T_s , and angular damping torques, $C_D R_D \dot{\theta}$, obtained from the physical

actuator/damper possibilities, were used to determine boom deployment times and velocities (Fig. 1); these analog results were used to improve the proposed actuator/damper reliability, performance, and weight in an iterative process. The results show that with an initial T_s of 26.4 lb-in./rad, $C_D R_D (\theta^{-1/2})$ of 716 lb-in.-s/rad, and reasonable steady and sticking friction coefficients (corresponding to Teflon sleeve bearings), the RTG can be fully deployed in about 70 s, with a residual energy on lock-open of less than 0.1 ft-lb. The generated friction coefficient has a high friction regime ($f_B = 0.2$) simulating sticking and roughness at low velocities (≤ 0.4 deg/s).

3. Current Actuator/Damper Design

When the physical actuator/damper possibilities were examined, four actuators (a split torsion cylinder, clock

spring, torsion spring, and torsion bar) and three dampers (rotating concentric tube viscous damper, fluid metering capillary damper, and pneumatic cylinder damper) were found to meet the torque/damping factor requirements. Of these, the torsion bar/concentric tube damper had the lowest weight (< 1.5 lb), because it could be used to replace the existing hinge axis structural member (Fig. 2), with end clamps and fittings incorporated into the boom hinge. The requisite T_s can be supplied by an aluminum torsion bar $\frac{3}{16}$ -in. diameter \times 17.5-in. active length. The damper typically consists of two 3-in. nominal tubes 18 in. long, separated by a 0.040-in. film of 1,000,000-cS silicone fluid (Fig. 3a). Number of tubes, nominal diameter, and fluid viscosity may be varied to achieve minimum weight, workable film thickness tolerances, and ease of fluid injection between the damper tubes (Fig. 3b).

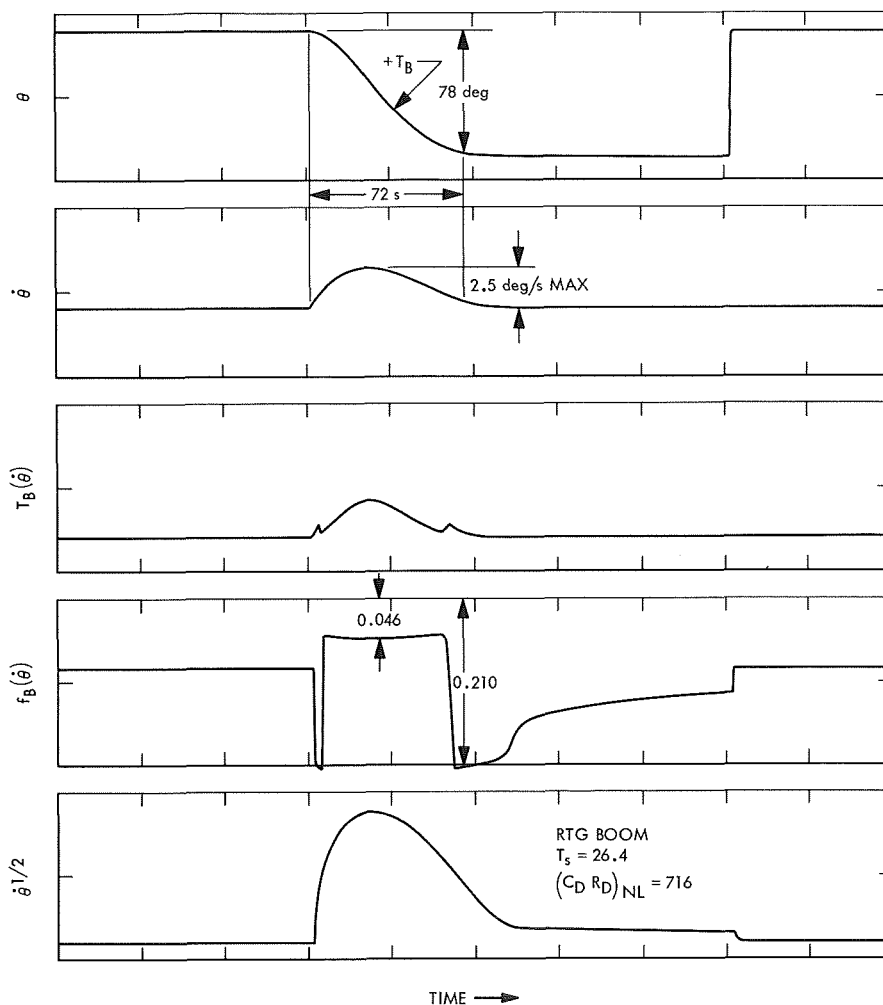


Fig. 1. Typical simultaneous analog computation of boom deployment angle θ , velocity $\dot{\theta}$, friction torque $T_B(\dot{\theta})$, friction coefficient $f_B(\dot{\theta})$, and nonlinear damping function $\dot{\theta}^{1/2}$ (not to same scale)

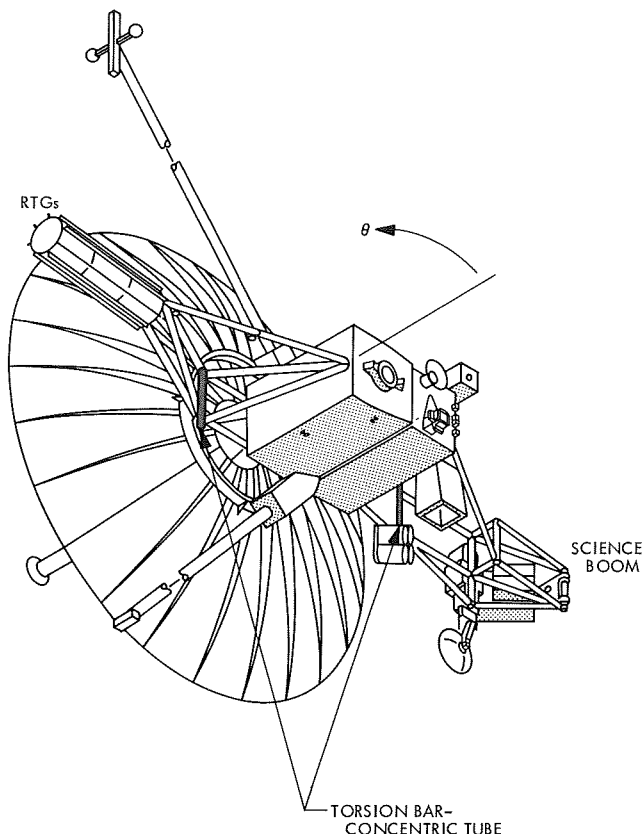


Fig. 2. Position of torsion bar-concentric tube actuator/dampers on TOPS spacecraft

The torsion bar actuator is particularly adaptable to changes in torque requirements while maintaining close tolerances on torsion constant T_s and on preload and end-point angles. To vary T_s , a bar with different diameter or effective length and a different set of collets can be installed, or for larger changes a material of higher elastic modulus may be specified.

The extremely low angular velocity of the RTG boom is necessitated by the requirements for low residual kinetic

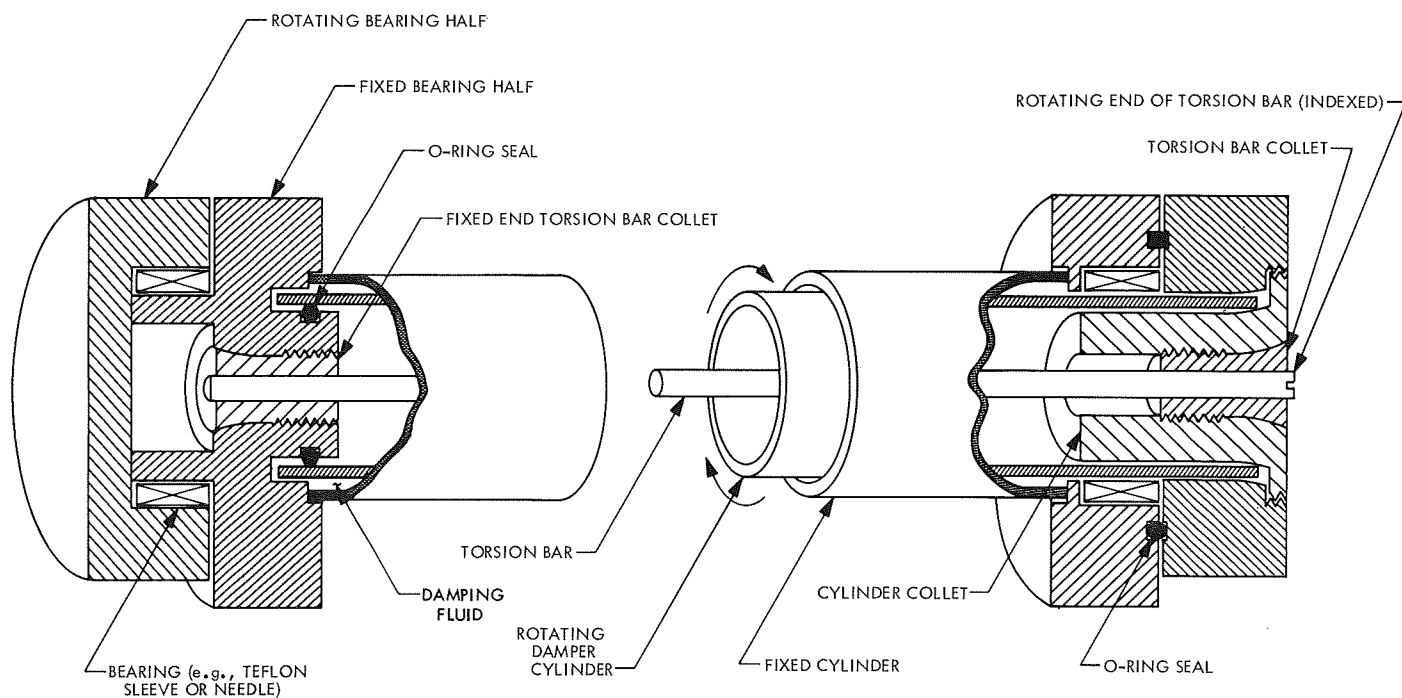
energy at lock-open and minimum perturbation of the spacecraft orientation. But it does require elimination of losses due to unaccounted-for bearing friction, hinge friction in the "elbow" truss member, and bending of electrical and fluid lines passing through boom hinges. Proposed solutions to this problem are to make lines and elbow hinges self-actuating by use of flexible strip hinges and spring-loaded electrical and fluid lines. Assuming these devices are practical, and that there are no friction losses other than in the hinge bearings, then the minimum-torque actuator should behave according to the free-space analog model. If moderate additional losses do exist (up to 10% of the actuator torque), the chief effect is to increase the opening time; possibility of the boom failing to deploy fully is minimized by locking the boom ($\theta = 0$) approximately 4 deg short of the torsion bar zero point.

Deployment tests on the ground impose losses greater than the minimum-torque actuator can tolerate, if the test is based on unsupported horizontal deployment, as is conventional in testing solar panel mechanisms. To eliminate the 1750-lb static cantilever load on the bearings, the RTG will be supported on an air bearing pad. Deviations of the hinge axis from vertical will be minimized by uncoupling the torsion bar and damper from the rotating outer boom, thus allowing it to act as its own leveling device.

4. Concluding Remarks

These proposals, as well as the initial assumptions for friction coefficients and weight estimates, will have to be verified on a full-scale mock-up of the RTG and boom, for which plans are presently being prepared. The RTG mock-up will consist of a 5.5-ft section of 10-in. steel pipe, and the boom of aluminum tubing, with the hinge axis vertically mounted on a wall. On conclusion of deployment tests with the mock-up, it will be used to determine electric and fluid line positioning and interfacing with the RTGs.

(a) DAMPERS A (1.5-in. OD) AND C (3-in. OD)



(b) DAMPERS B (1.5-in. OD) AND D (3-in. OD)

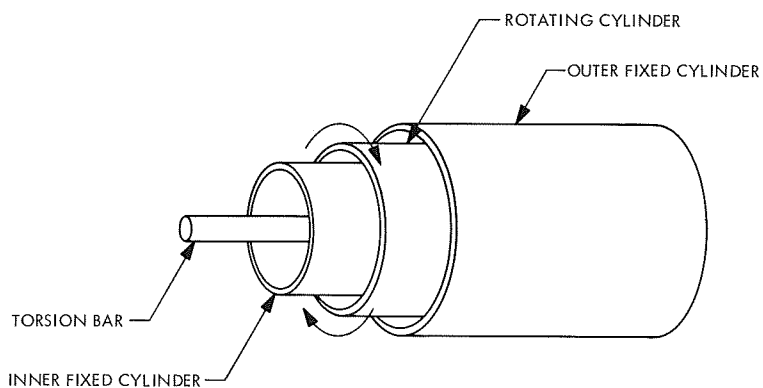


Fig. 3. Torsion bar actuator and concentric rotating cylinder damper configurations

XVII. Solid Propellant Engineering

PROPULSION DIVISION

A. A Statistical Model for Radiation Penetration in Composite Propellants, *R. L. Klaus*

1. Introduction

The role of radiation as a means of energy feedback to a burning solid propellant surface is not well understood. There seems to be evidence that the role of radiative heat transfer can be important, particularly in propellants which contain burning metal particles. The ignition characteristics of a propellant are certainly different if it is ignited by a radiant flux such as in the arc-image furnace or a more convective flux such as in a shock tube.

The difference in response of the burning surface to these two types of heat transfer is due to the fact that convective heat flux is delivered only to the surface, whereas radiant flux may penetrate the surface to a certain degree due to the fact that the propellant is not completely opaque to radiation. Oxidizer crystals are particularly transparent. Measurements (Ref. 1) have indicated that ammonium perchlorate (AP) can have absorptivities between 5×10^3 and 10^5 cm^{-1} at wavelengths between 1 and $0.34 \mu\text{m}$ and also that this transmissivity can be greatly altered by the presence of small amounts of impurities. Moreover, experimental measurements of temperature profiles in burning propellant (Ref. 2) have also indicated

a difference in temperature profile when the temperature was measured directly below an AP particle as compared to the temperature profile at a random location in the propellant. One plausible explanation of this difference in temperature profile is that it is caused by radiation penetration into the oxidizer crystal.

The detailed structure of a composite propellant is quite complex. Its properties are, therefore, probably best treated in a statistical way. That is, the propellant is imagined to consist of oxidizer crystals of varying sizes and shapes arranged at random in a matrix of binder. The gross properties of the propellant are then calculated by computing proper statistical averages of the individual constituents. The model suggested here has specific applicability to radiation penetration but the general approach might also be useful in other areas, such as heat transfer in the solid.

2. Development of the Mathematical Model

The mathematical model is based on the following assumptions:

- (1) The binder and any metallic constituents of the propellant are opaque to radiation, and the oxidizer crystals are translucent.

- (2) The oxidizer crystals are spheres which are randomly arranged throughout the propellant. In the first part of the analysis they are all considered to be of equal size, but a later generalization indicates how any size distribution may be treated.
- (3) The radiation flux is normal to the propellant surface. It is considered to be completely absorbed except where it strikes a protruding portion of an oxidizer crystal, in which case it penetrates as deeply as the crystal permits.
- (4) Due to the translucence of the crystal, the radiation may be partially absorbed as it passes through the crystal. In any event, it is assumed to be completely absorbed when it reaches the lower boundary of the crystal.

The first and crucial step in the development of this model is the statistical determination of the area available for radiation penetration as a function of depth of penetration into the propellant. We consider first the case in which all the oxidizer crystals are spheres of uniform diameter. The geometrical situation is illustrated in Fig. 1. The crystal may protrude above the surface, but this protrusion is ignored. It has no effect on the results for perfectly transparent crystals and only a secondary effect for translucent crystals. Thus the surface available for radiation penetration is that surface cut by the surface plane. The distance between the center plane of the sphere and the surface is defined to be x , and x is taken to be positive if the surface plane cuts the sphere in its lower hemisphere. At any distance below the surface, y , the area available for radiation penetration is the area within the

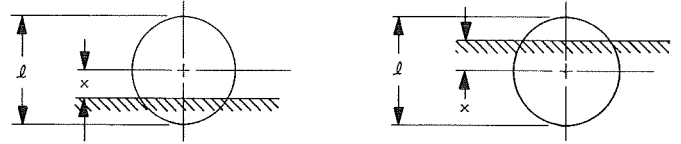


Fig. 1. Particle geometry and relationship to solid surface

sphere cut by a plane at a depth equal to y . Figure 1 also illustrates the case in which x is negative, i.e., more than half of the sphere is submerged in the propellant. In this case the area available for radiation penetration is taken to be equal to the area cut by the surface plane until a depth equal to $2x$ is reached, at which point the area is calculated as just described previously.

From this information the area available for radiation penetration due to the presence of a single crystal may be readily calculated. The available surface area A_s for a crystal whose center plane is at depth x , is given by

$$A_s(x) = \pi \left(\frac{\ell^2}{4} - x^2 \right)$$

This expression may be inverted to give

$$x = \pm \sqrt{\frac{\ell^2}{4} - \frac{A_s}{\pi}}$$

For $0 \leq x \leq \ell/2$, the area A available for radiation penetration at a distance y below the surface is given by

$$\begin{aligned} A(y) &= A_s - \pi y \left(y + 2 \sqrt{\frac{\ell^2}{4} - \frac{A_s}{\pi}} \right), & 0 \leq y \leq \frac{\ell}{2} - \sqrt{\frac{\ell^2}{4} - \frac{A_s}{\pi}} \\ A(y) &= 0, & y > \frac{\ell}{2} - \sqrt{\frac{\ell^2}{4} - \frac{A_s}{\pi}} \end{aligned} \quad (1)$$

For $-\ell/2 \leq x \leq 0$

$$\begin{aligned} A(y) &= A_s, & 0 \leq y < 2 \sqrt{\frac{\ell^2}{4} - \frac{A_s}{\pi}} \\ A(y) &= A_s - \pi y \left(y - 2 \sqrt{\frac{\ell^2}{4} - \frac{A_s}{\pi}} \right), & 2 \sqrt{\frac{\ell^2}{4} - \frac{A_s}{\pi}} \leq y < \frac{\ell}{2} + \sqrt{\frac{\ell^2}{4} - \frac{A_s}{\pi}} \\ A(y) &= 0, & y > \frac{\ell}{2} + \sqrt{\frac{\ell^2}{4} - \frac{A_s}{\pi}} \end{aligned} \quad (2)$$

A statistical analysis is necessary to properly calculate the average area available for radiation penetration at any depth y . We start with the basic assumption that if the surface plane cuts an oxidizer sphere, any value of x in the range $[-(\ell/2) \leq x \leq (\ell/2)]$ is equally likely. The formulas for the basic distribution function which will be used are

$$P(a \leq b \leq a + da) = f(b:a) da$$

$$P(b \leq a) = F(b:a) = \int_{-\infty}^a f(b:\xi) d\xi$$

That is, the probability that b lies between a and $a + da$ is equal to some function of a (the distribution function) multiplied by da . The probability that $b \leq a$ is the cumulative distribution function, designated by the symbol F . Since a certain event is accorded a probability of unity,

$$F(b:\infty) = 1 \quad (3)$$

With these formulas as a basis, the above statement about x may be expressed mathematically as

$$\begin{aligned} f(x:a) &= \frac{1}{\ell}, & -\frac{\ell}{2} \leq a \leq \frac{\ell}{2} \\ &= 0, & a \leq -\frac{\ell}{2}, \quad a \geq \frac{\ell}{2} \end{aligned}$$

The choice of a constant for the distribution function indicates that all values of x in the allowable range are equally likely, and the particular choice of the constant ensures that Eq. (3) is satisfied. From these facts it now follows that

$$F(|x|:a) = F(x:a) - F(x:-a) = \frac{2a}{\ell}$$

Statements can now be made about the distribution function of surface area.

$$f(A_s:\alpha) = \frac{1}{\sqrt{\frac{\ell^2}{4} - \frac{\alpha}{\pi}}}$$

The above equation is an important result in that it gives a formula for the probability distribution function of surface areas.

One now needs to know the number of particles per unit area of propellant surface which have a given area

exposed at the surface. This is calculated from the fact that the number of particles per unit area of propellant surface that are intersected by the surface plane in such a way that their surface area is between α and $\alpha + d\alpha$ is proportional to $f(A_s:\alpha)$. If this number is designated by $n(\alpha)$, the above statement may be expressed mathematically as follows:

$$n(\alpha) = kf(A_s:\alpha)$$

where k is a proportionality constant which may be determined from the fact that the sum of the exposed areas of all the particles must equal the total area available for radiation penetration at that particular depth. This area is designated R . Thus

$$\begin{aligned} \int_0^{\pi^{1/2}/4} \alpha n(\alpha) d\alpha &= \frac{k}{\pi\ell} \int_0^{\pi^{1/2}/4} \frac{\alpha d\alpha}{\sqrt{\frac{\ell^2}{4} - \frac{\alpha}{\pi}}} = \frac{k\pi\ell^2}{6} = R \\ k &= \frac{6R}{\pi\ell^2} \\ n(\alpha) &= \frac{6R}{\pi^2\ell^3} \frac{1}{\sqrt{\frac{\ell^2}{4} - \frac{\alpha}{\pi}}} \end{aligned} \quad (4)$$

It is now necessary to make an assumption about R . The surface area of oxidizer per unit area of propellant which is exposed when the propellant is cut at a random location is itself a random variable whose mean value is the volume fraction of the oxidizer crystals. For the purposes of the model, however, the approximation is made that R is a constant equal to the volume fraction of the oxidizer crystals.

It is now possible to calculate the total area available for radiation penetration. The fraction of nonreflected impinging radiation which penetrates at all is also equal to R . The area available for radiation penetration as a fraction of R will be designated by $S(y)$, which can be calculated according to the following formula:

$$\begin{aligned} S(y) &= \frac{1}{R} \left\{ \frac{1}{2} \int_0^{\pi^{1/2}/4} A_1(y,\alpha) n(\alpha) d\alpha \right. \\ &\quad \left. + \frac{1}{2} \int_0^{\pi^{1/2}/4} A_2(y,\alpha) n(\alpha) d\alpha \right\} \end{aligned}$$

where A_1 and A_2 refer to particle areas in which $x > 0$ and $x < 0$, respectively. Formulas for A_1 and A_2 are given by

Eqs. (1) and (2), respectively. These and Eq. (4) may be combined and the changes in the functions in different intervals taken into account by proper choice of limits of integration. One obtains

$$S(y) = \frac{3}{\pi^2 \ell^3} \left\{ \underbrace{\int_{\pi y(1-y)}^{\pi \ell^2/4} \frac{\left[\alpha - \pi y \left(y + 2 \sqrt{\frac{\ell^2}{4} - \frac{\alpha}{\pi}} \right) \right]}{\sqrt{\frac{\ell^2}{4} - \frac{\alpha}{\pi}}} d\alpha}_{0 \leq y \leq \frac{\ell}{2}} + \underbrace{\int_0^{\pi/4(\ell^2-y^2)} \frac{\alpha d\alpha}{\sqrt{\frac{\ell^2}{4} - \frac{\alpha}{\pi}}} }_{0 \leq y \leq \ell} \right. \\ \left. + \underbrace{\int_{\pi/4(\ell^2-y^2)}^{\pi \ell^2/4} \frac{\left[\alpha - \pi y \left(y - 2 \sqrt{\frac{\ell^2}{4} - \frac{\alpha}{\pi}} \right) \right]}{\sqrt{\frac{\ell^2}{4} - \frac{\alpha}{\pi}}} d\alpha}_{0 \leq y \leq \ell/2} + \underbrace{\int_{\pi(\ell^2-y^2)}^{\pi y(1-y)} \frac{\alpha d\alpha}{\sqrt{\frac{\ell^2}{4} - \frac{\alpha}{\pi}}} }_{\ell/2 \leq y \leq \ell} \right\}$$

which may finally be integrated to give

$$S(\xi) = (4\xi + 1)|2\xi - 1| \sqrt{\frac{1}{4} - \xi(1 - \xi)} \\ + \frac{1}{2} - \frac{\xi}{2} (15\xi^2 - 12\xi + 3) \quad (5)$$

where

$$\xi = \frac{y}{\ell}$$

This is the basic formula for the area available for radiation penetration as a function of penetration distance. Note that $S(0) = 1$ and $S(1) = 0$ as expected. Formulas for the first two derivatives are also of interest later and are therefore given here:

$$\frac{dS}{d\xi} = 6\xi \left[\frac{1}{4} - \xi(1 - \xi) \right]^{-1/2} (2\xi - 1) |2\xi - 1| \\ - \frac{1}{2} (45\xi^2 - 24\xi + 3) \quad (6)$$

$$\frac{d^2S}{d\xi^2} = 3|2\xi - 1| \left(8\xi^3 - 10\xi^2 + 4\xi - \frac{1}{2} \right) \left[\frac{1}{4} - \xi(1 - \xi) \right]^{-3/2} \\ - [45\xi - 12] \quad (7)$$

For perfectly opaque propellant and perfectly transparent oxidizer crystals, Eq. (5) is the formula for radiation transmitted at a given depth as a fraction of the total radiation transmitted, under the assumption that all oxidizer particles are spheres of diameter ℓ . The amount of radiation absorbed at depth y is equal to $-(dS/dy)$ which equals

$-(1/\ell) dS/d\xi$. This quantity may be readily calculated using Eq. (6).

3. Generalization to Translucent Crystals of Unequal Diameters

For the situation in which the oxidizer crystals are not completely transparent, the fraction of transmitted radiation which penetrates to level y , designated $I(y, \ell)$, is given by

$$I(y, \ell) = e^{-\beta y} S(y, \ell)$$

where β is the absorptivity. Thus the radiation absorbed is given by

$$J(y) = -\frac{dI}{dy} = e^{-\beta y} \left[\beta S(y, \ell) - \frac{dS}{dy}(y, \ell) \right]$$

or

$$\ell J(\xi) = e^{-(\beta \ell) \xi} \left[(\beta \ell) S(\xi) - \frac{dS}{d\xi}(\xi) \right] \quad (8)$$

Values of $\ell J(\xi)$ calculated according to this equation are plotted in Fig. 2 for various values of $(\beta \ell)$. The curve for $(\beta \ell) = 0$ applies to completely transparent oxidizer, while $(\beta \ell) = \infty$ applies to completely opaque crystals. Note that for fairly transparent crystals (low $\beta \ell$) the shape of the absorption curve is not exponential in shape at all. Radiation penetrates quite deeply into the surface but falls off sharply as the depth approaches the diameter of the crystals.

The next generalization is to the case in which there is a distribution of particles with various diameters. $\bar{J}(y)$,

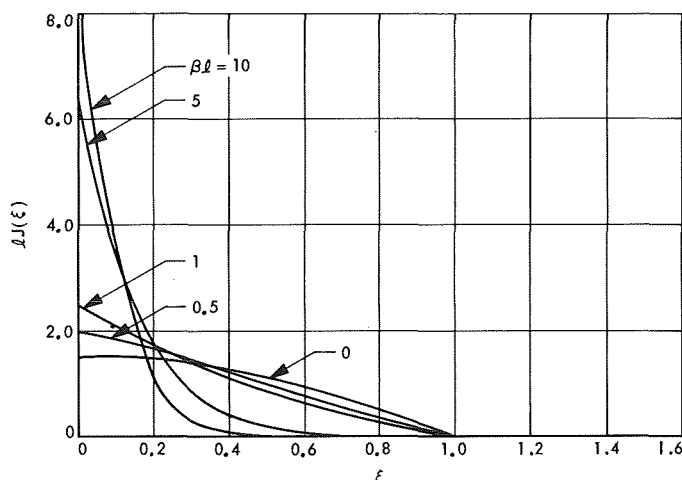


Fig. 2. $lJ(\xi)$ versus ξ for equally sized spheres

the overall fraction of transmitted radiation absorbed at depth y , is given by

$$\bar{J}(y) = \int_y^\infty J(y, \beta, \lambda) f(\ell; \lambda) d\lambda \quad (9)$$

$$F(\ell; \lambda) = \frac{1}{\sqrt{2\pi} \sigma N} \exp[-(\lambda - \bar{\ell})^2 / 2\sigma^2] - \exp[-(\bar{\ell}^2 / 2\sigma^2)], \quad 0 \leq \lambda \leq 2\bar{\ell}$$

$$= 0, \quad \lambda < 0, \quad \lambda > 2\bar{\ell}$$

where N is a normalizing factor, computed so as to give an area of unity under the probability distribution function. Its actual value can be calculated to be

$$N = \operatorname{erf}\left(\frac{\bar{\ell}}{\sqrt{2}\sigma}\right) - \sqrt{\frac{2}{\pi}} \frac{\bar{\ell}}{\sigma} \exp(-\bar{\ell}^2 / 2\sigma^2)$$

where erf is the error function defined as

$$\operatorname{erf}(x) = \frac{2}{\sqrt{\pi}} \int_0^x \exp(-\xi^2) d\xi$$

For values of $(\bar{\ell}^2 / 2\sigma^2)$, which are reasonably large, the distribution function does not differ very greatly from the normal distribution, but a function of this form eliminates the problems which would otherwise be present in the integration, because it goes exactly to zero at $\lambda = 0$. It should be noted that the distribution function also goes to zero at $\lambda = 2\bar{\ell}$. This distribution function may now be substituted back into Eq. (9). Since the J in that equation

In this equation, J is written as a function of three variables: penetration depth, absorptivity, and particle diameter. The integration is performed over all particle diameters greater than y , because there is clearly no contribution to \bar{J} due to particles with diameter less than y . This equation is useful where the distribution function of the particle diameters is known analytically, because a direct integration may be performed analytically to obtain \bar{J} .

One application of Eq. (9) is for a distribution of particle diameters which follows the normal distribution. That is,

$$F(\ell; \lambda) = \frac{1}{\sqrt{2\pi} \sigma} \exp[-(\lambda - \bar{\ell})^2 / 2\sigma^2]$$

where $\bar{\ell}$ is the mean particle diameter and σ is the standard deviation. However, this particle size distribution does not give exactly zero as its value for a particle diameter of zero. This fact would cause problems in the actual evaluation of the integrals. In order to circumvent this difficulty, we consider a particle size distribution as follows:

is expressed in terms of ξ , it is convenient to transform to ξ as a variable of integration. If in addition we define

$$\xi = \frac{y}{\ell}$$

the resulting equation becomes

$$\sigma J(\xi) = \frac{e^{-(\beta \bar{\ell}) \xi}}{\sqrt{2\pi} N} \int_{\xi/2}^1 \left[\beta \bar{\ell} \left(\frac{\xi}{\xi^2} \right) S(\xi) - \frac{1}{\xi} S'(\xi) \right]$$

$$\times \left[\exp\left[-\frac{\bar{\ell}^2}{2\sigma^2} \left(\frac{\xi}{\xi} - 1 \right)^2\right] - \exp\left(-\frac{\bar{\ell}^2}{2\sigma^2}\right) \right] d\xi$$

$$(\xi > 0) \quad (10)$$

This equation is not valid for $y = 0$ because the transformation to ξ as a variable of integration is not valid for $y = 0$. In the latter case, the change of variables is taken to be

$$\lambda = 2\bar{\ell}\xi$$

which, when substituted into Eq. (9) yields

$$\sigma J(0) = \sqrt{\frac{2}{\pi}} \frac{1}{N} \int_0^1 \left[\beta \bar{l} - \frac{3}{4\xi} \right] \left[\exp \left[-\frac{\bar{l}^2}{2\sigma^2} (2\xi - 1)^2 \right] - \exp \left(-\frac{\bar{l}^2}{2\sigma^2} \right) \right] d\xi \quad (11)$$

Notice that $\bar{J}(\xi)$ may be represented as a function of two parameters, $\beta \bar{l}$ and $\bar{l}^2/2\sigma^2$. A computer program has been written to perform the numerical integration indicated in Eqs. (10) and (11) and in Fig. 3, one typical plot is shown. Note again that the absorption does not fall off exponentially, particularly with relatively transparent particles.

In the most general case, however, particle diameters may not follow a normal distribution. If the cumulative number distribution function is available from experimental measurements, the radiation absorption may still be calculated by first integrating Eq. (9) by parts to obtain

$$\bar{J}(y) = J(y, \beta, a) - J(y, \beta, y) F(l:y) - \int_y^a \frac{\partial J}{\partial \lambda}(y, \beta, \lambda) F(l:\lambda) d\lambda$$

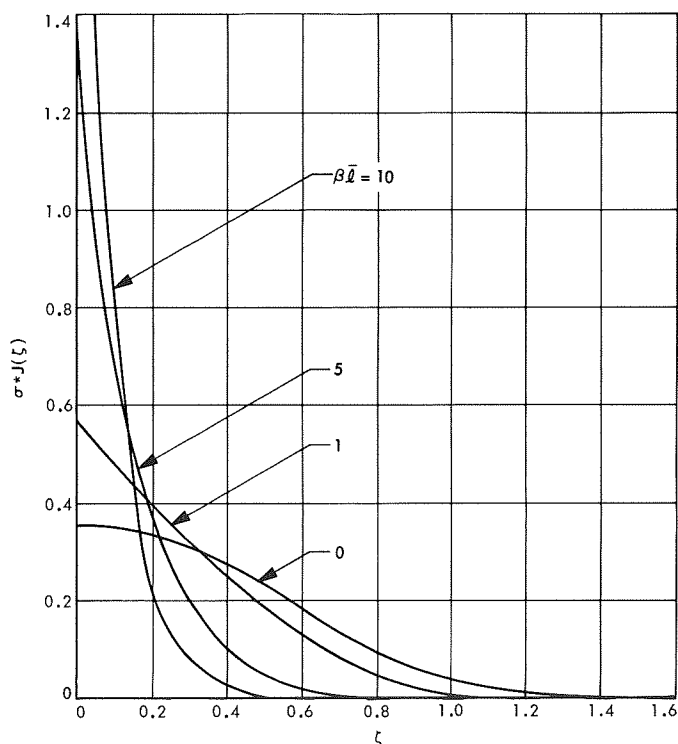


Fig. 3. $\sigma^* J(\xi)$ versus ξ for normal particle size distribution $\bar{l}^2/2\sigma^2 = 10$

where F is the cumulative number distribution function, and where a is the maximum particle diameter, i.e., $F(l:a) = 1$. Equation (8) may be used to reduce this equation to its final form:

$$\bar{J}(y) = e^{-\beta y} \left\{ \beta S\left(\frac{y}{a}\right) - \frac{1}{a} S'\left(\frac{y}{a}\right) + \int_y^a \frac{F(l:\lambda)}{\lambda^2} \left[(\beta y - 1) S'\left(\frac{y}{\lambda}\right) - \left(\frac{y}{\lambda}\right) S''\left(\frac{y}{\lambda}\right) \right] d\lambda \right\} \quad (12)$$

Equation (12) has been programmed for solution on a digital computer. Formulas for the indicated derivatives of S have already been derived in Eqs. (6) and (7).

Figure 4 shows a typical cumulative particle weight distribution for ammonium perchlorate of the type used in propellants. The measurements were made by the micromerograph technique (SPS 37-47, Vol. III, pp. 75-84). What is needed in Eq. (12), however, is the cumulative number distribution which can be calculated by the following formula:

$$F(l:b) = \frac{\int_0^b \left(\frac{6}{\pi \lambda^3} \right) w(l:\lambda) d\lambda}{\int_0^a \left(\frac{6}{\pi \lambda^3} \right) w(l:\lambda) d\lambda} = \frac{W(l:a)/3a^3 + \int_0^b \frac{W(l:\lambda)}{\lambda^4} d\lambda}{1/3a^3 + \int_0^a \frac{W(l:\lambda)}{\lambda^4} d\lambda}$$

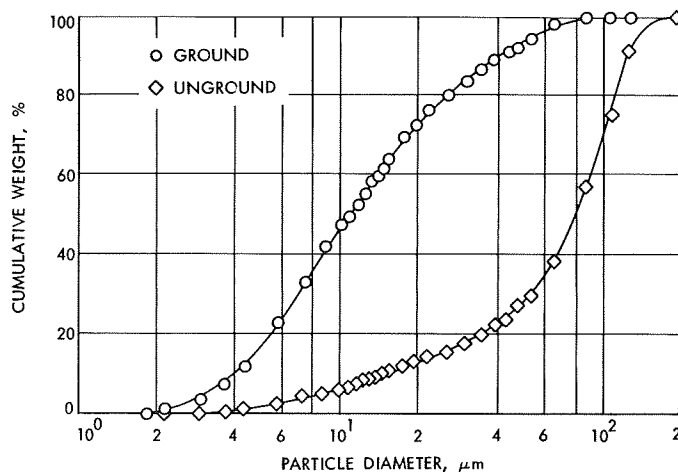


Fig. 4. Cumulative ammonium perchlorate particle weight distribution

where w is the weight fraction distribution function, and W is the cumulative weight fraction distribution function. The indicated integrations have also been incorporated into the computer program mentioned above. The calculated values of \bar{J} corresponding to one of the particle weight distributions shown in Fig. 4 is plotted in Fig. 5 for several values of absorptivity. It is significant to note that the depth to which radiation penetrates depends very strongly on the absorptivity. Two of the curves shown use experimentally determined absorptivities for different wavelengths, and it is clear that in one case radiation is essentially completely absorbed at the surface, whereas in the other, a significant fraction of the radiation penetrates to depths that are significant compared to the depth to which energy is transferred by conduction.

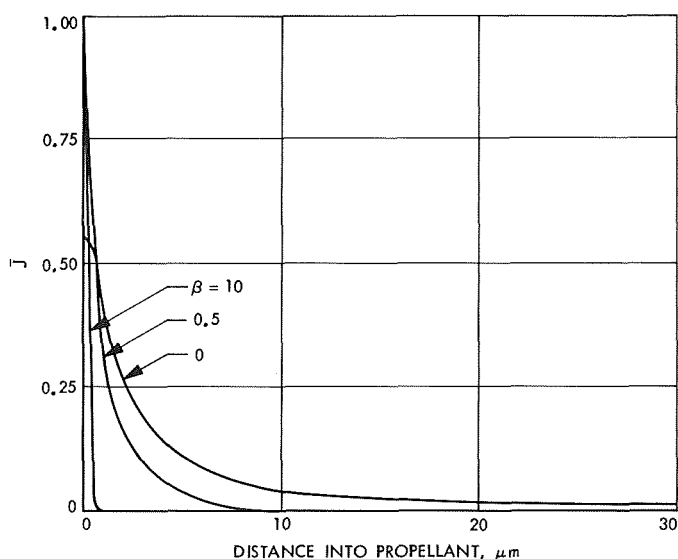


Fig. 5. Radiation absorption in ground ammonium perchlorate

The statistical approach presented here is admittedly only a first step in dealing with the properties of composite propellants, but it is felt that this is a direction which needs to be pursued if, in addition to radiation penetration, other phenomena such as heat conduction are to be properly treated mathematically.

References

1. Evans, M. W., Beyer, R. B., McCulley, L., "Initiation of Deflagration Waves at Surfaces of Ammonium Perchlorate-Copper Chromite-Carbon Pellets," *J. Chem. Phys.* Vol. 40, p 2431, May 1964.
2. *Combustion Tailoring Criteria for Solid Propellants*, LPC Report No. 835, Phase II. Lockheed Propulsion Co., Redlands, Calif.

B. Heat Conduction in a Semi-Infinite Solid With a Moving Surface With Application to Solid Propellant Rocket Combustion, R. L. Klaus

1. Introduction

The work reported here is a continuation of the work reported in SPS 37-57, Vol. III, pp. 133-139. In the former treatment, no account is taken of motion of the solid surface. This effect is included here. The inclusion of surface motion permits mathematical description of conduction into solids whose surface material is being removed by pyrolysis (such as propellants) or by ablation. The importance of this treatment to solid propellant combustion should be obvious, and since equations are presented for both the steady-state and transient cases, they are important in the study of both steady burning and transient burning such as occurs during ignition, extinction, and oscillatory combustion. As was done in the former treatment, account is taken of energy generation below the surface of the solid,¹ and three important types of boundary conditions. In addition, all integrals in the tabulated solutions are proper (wherever they are convergent) and converge to the correct values in the important limits. This is the first tabulation of this form and completeness which has appeared in the literature. After a discussion of the general governing equations, steady-state solutions are presented. Then a complete set of transient solutions are presented for the case of constant surface motion. Finally, a method is discussed which permits generalization of these equations to the case where the motion of the surface and convection coefficient both vary with time.

2. Mathematical Development

The basic governing equation for heat conduction in a semi-infinite solid with motion of the surface is

$$\frac{\partial^2 U}{\partial x^2} + \left(\frac{r}{\alpha}\right) \frac{\partial U}{\partial x} = \frac{1}{\alpha} \frac{\partial U}{\partial t} - g(x, t)$$

where U is the temperature, x the distance measured into the solid from the surface, t is time, r is the instantaneous rate of motion of the surface, and α is the thermal diffusivity which is taken to be constant.

¹The terms for energy generation below the surface of the solid can also be used to account for radiation penetration into the solid, which is important in the study of propellants ignited by a radiant energy flux.

There are three boundary conditions. The two which are always applied are

$$U(x, 0) = F(x)$$

$$U(\infty, t) = 0$$

The third can be one of three possible types:

Case 1 (specification of the surface temperature derivative):

$$\frac{\partial U}{\partial x}(0, t) = -h(t)$$

Case 2 (specification of the surface temperature):

$$U(0, t) = h(t)$$

Case 3 (convection at the surface):

$$U(0, t) - \frac{1}{p(t)} \frac{\partial U}{\partial x}(0, t) = h(t)$$

where $p(t)$ is the convective heat transfer coefficient divided by the thermal conductivity at the surface. This is considered given and may, as indicated, vary with time in the general case.

These equations may be directly solved under conditions which are far removed from steady state. For conditions which approach steady state it is convenient to make the following transformation:

$$U(x, t) = V(x) + W(x, t)$$

where V is the steady-state solution, and W is the extent to which the total solution differs from the steady solution. Each of these two functions may be considered separately. V satisfies the equation and boundary conditions shown in Table 1. The differential equation may be integrated twice and the integration constants evaluated from the boundary conditions. This procedure is straightforward and the resulting equations are also shown in Table 1.

When V and W are substituted into the equation for U , there results an equation for W (and boundary condi-

tions) of the following form:

$$\frac{\partial^2 W}{\partial x^2} + \left(\frac{r}{\alpha}\right) \frac{\partial W}{\partial x} = \frac{1}{\alpha} \frac{\partial W}{\partial t} - \left[\left(\frac{r}{\alpha}\right) - \left(\frac{r}{\alpha}\right)_\infty\right] \frac{dV}{dx} - [g(x, t) - g(x, \infty)]$$

$$W(x, 0) = F(x)$$

$$W(\infty, t) = 0$$

Case 1

$$\frac{\partial W}{\partial x}(0, t) = -[h(t) - h(\infty)]$$

Case 2

$$W(0, t) = h(t) - h(\infty)$$

Case 3

$$W(0, t) - \frac{1}{p(t)} \frac{\partial W}{\partial x}(0, t) = \left[\frac{1}{p(t)} - \frac{1}{p(\infty)}\right] \frac{dV}{dx}(0) + [h(t) - h(\infty)]$$

These equations are actually of the same form as the original equations for U . The basic form is shown in Table 2 with the meanings of the functions also specified depending on whether T refers to U or W . Thus the solutions which are also tabulated there may be used either in solutions far removed from or approaching the steady state. As can be seen from the above boundary conditions, the parts of the solution due to the surface boundary condition and heat generation term vanish as time becomes large. And as can be seen from the tabulated solutions, the part of the solution due to the initial distribution becomes vanishingly small for large time. Thus the solution of the equation in W approaches zero for large time, ensuring that the total solution approaches steady state for large time. From a numerical standpoint it is advantageous to use the equation for W at large times, due to the fact that the solutions become a small fraction of the total solution at a time when the errors in the solution tend to grow because the integrations are performed over long intervals in time. Thus the total error in the solutions is kept to a minimum. At small values of time, on the other hand, it is more convenient to solve the equation in U .

Table 1. Solutions of the steady state heat conduction equation in a semi-infinite solid with heat generation

Basic equations	
$\frac{d^2V}{dx^2} + \left(\frac{r_\infty}{\alpha}\right) \frac{dV}{dx} = -g_\infty(x)$	
Boundary condition which is always applied: $V(\infty) = 0$	
Additional boundary condition:	
Case 1 (surface temperature derivative specification)	$\frac{dV}{dx}(0) = -h_\infty$
Case 2 (surface temperature specification)	$V(0) = h_\infty$
Case 3 (convection)	$V(0) - \frac{1}{p_\infty} \frac{dV}{dx}(0) = h_\infty$
Solutions	
Case 1	
$V = \frac{\alpha}{r_\infty} \left[h_\infty \exp\left[-\left(\frac{r_\infty x}{\alpha}\right)\right] + \int_0^\infty g_\infty(\xi) d\xi \right] - \int_0^x \exp\left(-\left[\frac{r_\infty}{\alpha}(x-\zeta)\right]\right) \int_0^\zeta g_\infty(\xi) d\xi d\zeta$	
$\frac{dV}{dx} = -h_\infty \exp\left[-\left(\frac{r_\infty x}{\alpha}\right)\right] - \int_0^x g_\infty(\xi) d\xi + \left(\frac{r_\infty}{\alpha}\right) \int_0^x \exp\left(-\left[\frac{r_\infty}{\alpha}(x-\zeta)\right]\right) \int_0^\zeta g_\infty(\xi) d\xi d\zeta$	
Case 2	
$V = h_\infty \exp\left[-\left(\frac{r_\infty x}{\alpha}\right)\right] + \left(\frac{\alpha}{r_\infty}\right) \left[1 - \exp\left[-\left(\frac{r_\infty x}{\alpha}\right)\right] \right] \int_0^\infty g_\infty(\xi) d\xi$	
$- \int_0^x \exp\left(-\left[\frac{r_\infty}{\alpha}(x-\zeta)\right]\right) \int_0^\zeta g_\infty(\xi) d\xi d\zeta$	
$\frac{dV}{dx} = -\left(\frac{r_\infty}{\alpha}\right) h_\infty \exp\left[-\left(\frac{r_\infty x}{\alpha}\right)\right] + \exp\left[-\left(\frac{r_\infty x}{\alpha}\right)\right] \int_0^\infty g_\infty(\xi) d\xi - \int_0^x g_\infty(\xi) d\xi$	
$+ \left(\frac{r_\infty}{\alpha}\right) \int_0^x \exp\left(-\left[\frac{r_\infty}{\alpha}(x-\zeta)\right]\right) \int_0^\zeta g_\infty(\xi) d\xi d\zeta$	
Case 3	
$V = \frac{\left[h_\infty - \left(\frac{\alpha}{r_\infty}\right) \int_0^\infty g_\infty(\xi) d\xi \right] \exp\left[-\left(\frac{r_\infty x}{\alpha}\right)\right] + \left(\frac{\alpha}{r_\infty}\right) \int_0^\infty g_\infty(\xi) d\xi - \int_0^x \exp\left[-\frac{r_\infty}{\alpha}(x-\zeta)\right] \int_0^\zeta g_\infty(\xi) d\xi d\zeta}{\left[1 + \frac{r_\infty}{\alpha p_\infty} \right]}$	
$\frac{dV}{dx} = \frac{\left[\int_0^\infty g_\infty(\xi) d\xi - \left(\frac{r_\infty}{\alpha}\right) h_\infty \right] \exp\left[-\left(\frac{r_\infty}{\alpha}x\right)\right]}{\left[1 + \frac{r_\infty}{\alpha p_\infty} \right]}$	
$+ \left(\frac{r_\infty}{\alpha}\right) \int_0^x \exp\left(-\left[\frac{r_\infty}{\alpha}(x-\zeta)\right]\right) \int_0^\zeta g_\infty(\xi) d\xi d\zeta - \int_0^x g_\infty(\xi) d\xi$	

Table 2. General solutions of the heat conduction equation in a semi-infinite solid with heat generation

Basic equations	
$\frac{\partial^2 T}{\partial x^2} + \left(\frac{r}{\alpha}\right) \frac{\partial T}{\partial x} = \frac{1}{\alpha} \frac{\partial T}{\partial t} - G(x, t)$	
Boundary conditions which are always applied $T(x, 0) = F(x)$ $T(\infty, t) = 0$	
Additional boundary condition:	
Case 1 (Surface temperature derivative specification)	$\frac{\partial T}{\partial x}(0, t) = -H(t)$
Case 2 (Surface temperature specification)	$T(0, t) = H(t)$
Case 3 (Convection)	$T(0, t) - \frac{1}{p} \frac{\partial T}{\partial x}(0, t) = H(t)$
Meaning of the functions	
For $T = U$:	$G(x, t) = g(x, t)$ $H(t) = h(t)$
For $T = W$:	$G(x, t) = [g(x, t) - g(x, \infty)] + \left[\frac{r - r_\infty}{\alpha}\right] \frac{dV}{dx}(x)$ $H(t) = h(t) - h(\infty) \quad (\text{Cases 1 and 2})$ $H(t) = h(t) - h(\infty) + \left[\frac{1}{p(t)} - \frac{1}{p(\infty)}\right] \frac{dV}{dx}(0) \quad (\text{Case 3})$
Solutions	
Case 1	$T = T_{s1} + T_{i1} + T_{\theta1}$
Case 2	$T = T_{s2} + T_{i2} + T_{\theta2}$
Case 3	$T = T_{s3} + T_{i3} + T_{\theta3}$
Details of the functions	
T_{s1} (surface temperature derivative specification)	
$T_{s1} = \left[\frac{r}{2} + \left(\frac{2\alpha}{r} \right) \frac{\partial}{\partial t} \right] \int_0^t H(t - \lambda) \left\{ \exp \left[\frac{-r}{2\alpha} \left(x + \frac{r\lambda}{2} \right) \right] \operatorname{erfc} \left(\frac{x}{2\sqrt{\alpha\lambda}} \right) - \operatorname{erfc} \left[\frac{x + r\lambda}{2\sqrt{\alpha\lambda}} \right] \right\} d\lambda$ $\frac{\partial T_{s1}}{\partial x} = - \left[\left(\frac{r^2}{4\alpha} \right) + \frac{\partial}{\partial t} \right] \int_0^t H(t - \lambda) \left\{ \exp \left[-\frac{r}{2\alpha} \left(x + \frac{r\lambda}{2} \right) \right] \operatorname{erfc} \left(\frac{x}{2\sqrt{\alpha\lambda}} \right) \right\} d\lambda$	

Table 2 (contd)

$T_{s1}(0, t) = \left[\frac{r}{2} + \left(\frac{2\alpha}{r} \right) \frac{\partial}{\partial t} \right] \int_0^t H(t - \lambda) \left\{ \exp \left[\frac{-r^2 \lambda}{4\alpha} \right] - \operatorname{erfc} \left[\frac{r}{2} \sqrt{\frac{\lambda}{\alpha}} \right] \right\} d\lambda$ $\frac{\partial T_{s1}}{\partial x}(0, t) = -H(t) \quad t > 0$	$T_{s1}(x, 0) = 0$ $\frac{\partial T_{s1}}{\partial x}(x, 0) = 0$
<p>For $r = 0$</p> $T_{s1} = \frac{\partial}{\partial t} \int_0^t H(t - \lambda) \left\{ 2 \sqrt{\frac{\alpha \lambda}{\pi}} \exp \left(\frac{-x^2}{4\alpha \lambda} \right) - x \operatorname{erfc} \left(\frac{x}{2 \sqrt{\alpha \lambda}} \right) \right\} d\lambda$ $\frac{\partial T_{s1}}{\partial x} = - \frac{\partial}{\partial t} \int_0^t H(t - \lambda) \operatorname{erfc} \left(\frac{x}{2 \sqrt{\alpha \lambda}} \right) d\lambda$ $T_{s1}(0, t) = 2 \frac{\partial}{\partial t} \int_0^t H(t - \lambda) \operatorname{erfc} \left(\frac{x}{2 \sqrt{\alpha \lambda}} \right) d\lambda$	
<p>T_{i1} (initial temperature distribution)</p>	
$T_{i1} = \frac{1}{2} \int_0^\infty F(\xi) \left\{ \frac{1}{\sqrt{\pi \alpha t}} \exp \left[\frac{-(x - \xi + rt)^2}{4\alpha t} \right] \left[1 + \exp \left(\frac{-x\xi}{\alpha t} \right) \right] - \left(\frac{r}{\alpha} \right) \exp \left(\frac{r\xi}{\alpha} \right) \operatorname{erfc} \left[\frac{rt + x + \xi}{2 \sqrt{\alpha t}} \right] \right\} d\xi$ $\frac{\partial T_{i1}}{\partial x} = \frac{-1}{4 \sqrt{\pi (\alpha t)^3}} \int_0^\infty F(\xi) \exp \left[\frac{-(x - \xi + rt)^2}{4\alpha t} \right] \left\{ [x - \xi + rt] + [x + \xi - rt] \exp \left(\frac{-x\xi}{\alpha t} \right) \right\} d\xi$	
$T_{i1} = F(x + rt) \left\{ 1 + \exp \left[\frac{-x}{\alpha} \left(\frac{x}{t} + r \right) \right] \right\} - F(0) \operatorname{erfc} \left[\frac{(x + rt)}{2 \sqrt{\alpha t}} \right]$ $- \frac{1}{2} \left\{ \int_0^{x+rt} \frac{d}{d\xi} \left\{ F(\xi) \left[1 + \exp \left(\frac{-x\xi}{\alpha t} \right) \right] \right\} \operatorname{erfc} \left[\frac{x - \xi + rt}{2 \sqrt{\alpha t}} \right] d\xi \right.$ $- \int_{x+rt}^\infty \frac{d}{d\xi} \left\{ F(\xi) \left[1 + \exp \left(\frac{-x\xi}{\alpha t} \right) \right] \right\} \operatorname{erfc} \left[\frac{\xi - x - rt}{2 \sqrt{\alpha t}} \right] d\xi$ $\left. + \left(\frac{r}{\alpha} \right) \int_0^\infty F(\xi) \exp \left(\frac{r\xi}{\alpha} \right) \operatorname{erfc} \left[\frac{x + \xi + rt}{2 \sqrt{\alpha t}} \right] d\xi \right\}$ $\frac{\partial T_{i1}}{\partial x} = F'(x + rt) \left\{ 1 - \exp \left[\frac{-x}{\alpha} \left(\frac{x}{t} + r \right) \right] \right\} - \frac{1}{2} \left\{ \int_0^{x+rt} \frac{d}{d\xi} \left\{ F'(\xi) \left[1 - \exp \left(\frac{-x\xi}{\alpha t} \right) \right] \right\} \right.$ $\left. \times \operatorname{erfc} \left[\frac{x - \xi + rt}{2 \sqrt{\alpha t}} \right] d\xi - \int_{x+rt}^\infty \frac{d}{d\xi} \left\{ F'(\xi) \left[1 - \exp \left(\frac{-x\xi}{\alpha t} \right) \right] \right\} \operatorname{erfc} \left[\frac{\xi - x - rt}{2 \sqrt{\alpha t}} \right] d\xi \right\}$	
$T_{i1}(0, t) = 2 F(rt) - F(0) \operatorname{erfc} \left(\frac{r}{2} \sqrt{\frac{t}{\alpha}} \right) - \int_0^{rt} F'(\xi) \operatorname{erfc} \left(\frac{rt - \xi}{2 \sqrt{\alpha t}} \right) d\xi$ $+ \int_{rt}^\infty F'(\xi) \operatorname{erfc} \left[\frac{\xi - rt}{2 \sqrt{\alpha t}} \right] d\xi - \left(\frac{r}{2\alpha} \right) \int_0^\infty F(\xi) \exp \left(\frac{r\xi}{\alpha} \right) \operatorname{erfc} \left(\frac{\xi + rt}{2 \sqrt{\alpha t}} \right) d\xi$ $\frac{\partial T_{i1}}{\partial x}(0, t) = 0 \quad (t > 0)$	$T_{i1}(x, 0) = F(x)$ $\frac{\partial T_{i1}}{\partial x}(x, 0) = F'(x)$

Table 2 (contd)

T_{g1} (heat generation or radiation penetration)	
$T_{g1} = \frac{1}{2} \int_0^t \int_0^\infty G(\xi, t - \lambda) \left\{ \sqrt{\frac{\alpha}{\lambda\pi}} \left[1 + \exp\left(\frac{-x\xi}{\alpha\lambda}\right) \right] \exp\left[\frac{-(x - \xi + r\lambda)^2}{4\alpha\lambda}\right] \right.$ $\left. - r \exp\left(\frac{r\xi}{\alpha}\right) \operatorname{erfc}\left[\frac{\xi + x + r\lambda}{2\sqrt{\alpha\lambda}}\right] \right\} d\xi d\lambda$ $\frac{\partial T_{g1}}{\partial x} = -\frac{1}{4} \int_0^t \int_0^\infty \frac{G(\xi, t - \lambda)}{\sqrt{\pi(\alpha\lambda)^3}} \exp\left[\frac{-(x - \xi + r\lambda)^2}{4\alpha\lambda}\right] \left\{ [x - \xi + r\lambda] \right.$ $\left. + [x + \xi - r\lambda] \exp\left(\frac{-x\xi}{\alpha\lambda}\right) \right\} d\xi d\lambda$	
$T_{g1}(0, t) = \int_0^t \int_0^\infty G(\xi, t - \lambda) \left\{ \sqrt{\frac{\alpha}{\lambda\pi}} \exp\left[\frac{-(r\lambda - \xi)^2}{4\alpha\lambda}\right] \right.$ $\left. - r \exp\left(\frac{r\xi}{\alpha}\right) \operatorname{erfc}\left[\frac{\xi + r\lambda}{2\sqrt{\alpha\lambda}}\right] \right\} d\xi d\lambda$ $\frac{\partial T_{g1}}{\partial x}(0, t) = 0$	$T_{g1}(x, 0) = 0$ $\frac{\partial T_{g1}}{\partial x}(x, 0) = 0$
T_{s2} (surface temperature specification)	
$T_{s2} = \left[\frac{r^2}{4\alpha} + \frac{\partial}{\partial t} \right] \int_0^t H(t - \lambda) \exp\left[\frac{-r}{2\alpha} \left(x + \frac{r\lambda}{2} \right) \right] \operatorname{erfc}\left[\frac{x}{2\sqrt{\alpha\lambda}}\right] d\lambda$ $\frac{\partial T_{s2}}{\partial x} = -\left[\frac{r^2}{4\alpha} + \frac{\partial}{\partial t} \right] \int_0^t H(t - \lambda) \exp\left[-\frac{r}{2\alpha} \left(x + \frac{r\lambda}{2} \right) \right] \left\{ \left(\frac{r}{2\alpha} \right) \operatorname{erfc}\left[\frac{x}{2\sqrt{\alpha\lambda}}\right] + \frac{1}{\sqrt{\alpha\pi\lambda}} \exp\left[\frac{-x^2}{4\alpha\lambda}\right] \right\} d\lambda$	
$T_{s2}(0, t) = H(t) \quad (t > 0)$ $\frac{\partial T_{s2}}{\partial x}(0, t) = -\left[\frac{r^2}{4\alpha} + \frac{\partial}{\partial t} \right] \left\{ 2\sqrt{\frac{t}{\pi\alpha}} H(0) \exp\left[\frac{-r^2 t}{4\alpha}\right] + \int_0^t \left\{ \left(\frac{r}{2\alpha} \right) \right. \right.$ $\left. \left. + \frac{r^2}{2} \sqrt{\frac{\lambda}{\pi\alpha^3}} \right\} H(t - \lambda) + 2\sqrt{\frac{t}{\pi\alpha}} H'(t - \lambda) \right\} \exp\left[\frac{-r^2 \lambda}{4\alpha}\right] d\lambda \right\}$	$T_{s2}(x, 0) = 0$ $\frac{\partial T_{s2}}{\partial x}(x, 0) = 0$
T_{i2} (initial temperature distribution)	
$T_{i2} = \frac{1}{2\sqrt{\pi\alpha t}} \int_0^\infty F(\xi) \exp\left[\frac{-(x - \xi + rt)^2}{4\alpha t}\right] \left\{ 1 - \exp\left(\frac{-x\xi}{\alpha t}\right) \right\} d\xi$ $\frac{\partial T_{i2}}{\partial x} = \frac{-1}{4\sqrt{\pi(\alpha t)^3}} \int_0^\infty F(\xi) \exp\left[\frac{-(x - \xi + rt)^2}{4\alpha t}\right] \left\{ [x - \xi + rt] - [x + \xi + rt] \exp\left(\frac{-x\xi}{\alpha t}\right) \right\} d\xi$	

Table 2 (contd)

$T_{i2} = F(x + rt) \left\{ 1 - \exp \left[\frac{-x}{\alpha} \left(\frac{x}{t} + r \right) \right] \right\} - \frac{1}{2} \left\{ \int_0^{x+rt} \frac{d}{d\xi} \left\{ F(\xi) \left[1 - \exp \left(\frac{-x\xi}{\alpha t} \right) \right] \right\} \right.$ $\times \operatorname{erfc} \left[\frac{x - \xi + rt}{2\sqrt{\alpha t}} \right] d\xi - \int_{x+rt}^{\infty} \frac{d}{d\xi} \left\{ F(\xi) \left[1 - \exp \left(\frac{-x\xi}{\alpha t} \right) \right] \right\} \operatorname{erfc} \left[\frac{\xi - x - rt}{2\sqrt{\alpha t}} \right] d\xi \left. \right\}$ $\frac{\partial T_{i2}}{\partial x} = F'(x + rt) - F'(0) \operatorname{erfc} \left(\frac{x + rt}{2\sqrt{\alpha t}} \right) - \left(\frac{r}{2\alpha} \right) F(0) \operatorname{erfc} \left(\frac{x + rt}{2\sqrt{\alpha t}} \right)$ $- \frac{1}{2} \left\{ \int_0^{x+rt} F''(\xi) \operatorname{erfc} \left[\frac{(x - \xi + rt)}{2\sqrt{\alpha t}} \right] d\xi - \int_{x+rt}^{\infty} F''(\xi) \operatorname{erfc} \left[\frac{\xi - x - rt}{2\sqrt{\alpha t}} \right] d\xi \right.$ $\left. + \int_0^{\infty} \frac{d^2}{d\xi^2} \left[F(\xi) \exp \left(\frac{r\xi}{\alpha} \right) \right] \operatorname{erfc} \left[\frac{x + \xi + rt}{2\sqrt{\alpha t}} \right] d\xi \right\}$	
$T_{i2}(0, t) = 0$ $\frac{\partial T_{i2}}{\partial x}(0, t) = F'(rt) - \left[F'(0) + \left(\frac{r}{2\alpha} \right) F(0) \right] \operatorname{erfc} \left(\frac{r}{2} \sqrt{\frac{t}{\alpha}} \right)$ $- \frac{1}{2} \left\{ \int_0^{rt} F''(\xi) \operatorname{erfc} \left[\frac{rt - \xi}{2\sqrt{\alpha t}} \right] d\xi - \int_{rt}^{\infty} F''(\xi) \operatorname{erfc} \left[\frac{\xi - rt}{2\sqrt{\alpha t}} \right] d\xi \right.$ $\left. + \int_0^{\infty} \frac{d^2}{d\xi^2} \left[F(\xi) \exp \left(\frac{r\xi}{\alpha} \right) \right] \operatorname{erfc} \left[\frac{\xi + rt}{2\sqrt{\alpha t}} \right] d\xi \right\}$	$T_{i2}(x, 0) = F(x)$ $\frac{\partial T_{i2}}{\partial x}(x, 0) = F'(x)$
T_{g2} (heat generation or radiation penetration)	
$T_{g2} = \frac{1}{2} \int_0^t \int_0^{\infty} G(\xi, t - \lambda) \sqrt{\frac{\alpha}{\pi \lambda}} \exp \left[\frac{-(x - \xi + r\lambda)^2}{4\alpha \lambda} \right] \left[1 - \exp \left(\frac{-x\xi}{\alpha \lambda} \right) \right] d\xi d\lambda$ $\frac{\partial T_{g2}}{\partial x} = - \frac{1}{4} \int_0^t \int_0^{\infty} \frac{G(\xi, t - \lambda)}{\sqrt{\pi \alpha \lambda^3}} \exp \left[\frac{-(x - \xi + r\lambda)^2}{4\alpha \lambda} \right] \left\{ [x - \xi + r\lambda] - [x + \xi + r\lambda] \exp \left(\frac{-x\xi}{\alpha \lambda} \right) \right\} d\xi d\lambda$	
$T_{g2}(0, t) = 0$ $\frac{\partial T_{g2}}{\partial x}(0, t) = - \frac{1}{2} \int_0^t \int_0^{\infty} \frac{G(\xi, t - \lambda)}{\sqrt{\pi \alpha \lambda^3}} \xi \exp \left[\frac{-(r\lambda - \xi)^2}{4\alpha \lambda} \right]$	$T_{g2}(x, 0) = 0$ $\frac{\partial T_{g2}}{\partial x}(x, 0) = 0$

Table 2 (contd)

T_{ss} (convection boundary condition)	
$T_{ss} = \left[\frac{\frac{r^2}{4\alpha} + \frac{\partial}{\partial t}}{1 + \frac{r}{2\alpha p}} \right] \int_0^t H(t-\lambda) \left\{ \exp \left[-\frac{r}{2\alpha} \left(x + \frac{r\lambda}{2} \right) \right] \operatorname{erfc} \left(\frac{x}{2\sqrt{\alpha\lambda}} \right) \right.$ $\left. - \exp [p(x + r\lambda + \alpha p\lambda)] \operatorname{erfc} \left[\frac{x + r\lambda + 2\alpha p\lambda}{2\sqrt{\alpha\lambda}} \right] \right\} d\lambda$ $\frac{\partial T_{ss}}{\partial x} = - \left[\frac{\frac{r^2}{4\alpha} + \frac{\partial}{\partial t}}{1 + \frac{r}{2\alpha p}} \right] \int_0^t H(t-\lambda) \left\{ \left(\frac{r}{2\alpha} \right) \exp \left[-\frac{r}{2\alpha} \left(x + \frac{r\lambda}{2} \right) \right] \operatorname{erfc} \left(\frac{x}{2\sqrt{\alpha\lambda}} \right) \right.$ $\left. + p \exp [p(x + r\lambda + \alpha p\lambda)] \operatorname{erfc} \left[\frac{x + r\lambda + 2\alpha p\lambda}{2\sqrt{\alpha\lambda}} \right] \right\} d\lambda$	
$T_{ss}(0, t) = \left[\frac{\frac{r^2}{4\alpha} + \frac{\partial}{\partial t}}{1 + \frac{r}{2\alpha p}} \right] \int_0^t H(t-\lambda) \left\{ \exp \left(\frac{-r^2\lambda}{4\alpha} \right) \right.$ $\left. - \exp [p\lambda(\alpha p + r)] \operatorname{erfc} \left[\frac{1}{2} \sqrt{\frac{\lambda}{\alpha}} (r + 2\alpha p) \right] \right\} d\lambda$ $\frac{\partial T_{ss}}{\partial x}(0, t) = - \left[\frac{\frac{r^2}{4\alpha} + \frac{\partial}{\partial t}}{1 + \frac{r}{2\alpha p}} \right] \int_0^t H(t-\lambda) \left\{ \left(\frac{r}{2\alpha} \right) \exp \left(\frac{-r^2\lambda}{4\alpha} \right) \right.$ $\left. + p \exp [p\lambda(\alpha p + r)] \operatorname{erfc} \left[\frac{1}{2} \sqrt{\frac{\lambda}{\alpha}} (r + 2\alpha p) \right] \right\} d\lambda$	$T_{ss}(x, 0) = 0$ $\frac{\partial T_{ss}}{\partial x}(x, 0) = 0$
T_{i3} (initial temperature distribution)	
$T_{i3} = \frac{1}{2} \int_0^\infty F(\xi) \left\{ \frac{1}{\sqrt{\pi\alpha t}} \exp \left[\frac{-(x - \xi + rt)^2}{4\alpha t} \right] \left[1 + \exp \left(\frac{-x\xi}{\alpha t} \right) \right] \right.$ $\left. - \left(\frac{r}{\alpha} + 2p \right) \exp \left[\frac{r\xi}{\alpha} + p(x + \xi + rt + p\alpha t) \right] \operatorname{erfc} \left[\frac{x + \xi + rt + 2p\alpha t}{2\sqrt{\alpha t}} \right] \right\} d\xi$ $\frac{\partial T_{i3}}{\partial x} = - \frac{1}{2} \int_0^\infty F(\xi) \left\{ \frac{1}{2\sqrt{\pi(\alpha t)^3}} \exp \left[\frac{-(rt + x - \xi)^2}{4\alpha t} \right] \left\{ (x - \xi + rt) \right. \right.$ $\left. + (x + \xi - rt - 4p\alpha t) \exp \left(\frac{-x\xi}{\alpha t} \right) \right\} + p \left(\frac{r}{\alpha} + 2p \right) \exp \left[\frac{r\xi}{\alpha} + p(x + \xi + rt + p\alpha t) \right]$ $\times \operatorname{erfc} \left[\frac{x + \xi + rt + 2p\alpha t}{2\sqrt{\alpha t}} \right] \left. \right\} d\xi$	

Table 2 (contd)

$$\begin{aligned}
T_{i3} = & F(x+rt) \left\{ 1 + \exp \left[\frac{-x}{\alpha} \left(\frac{x}{t} + r \right) \right] \right\} - F(0) \operatorname{erfc} \left[\frac{x+rt}{2\sqrt{\alpha t}} \right] \\
& - \frac{1}{2} \left\{ \int_0^{x+rt} \frac{d}{d\xi} \left\{ F(\xi) \left[1 + \exp \left(\frac{-x\xi}{\alpha t} \right) \right] \right\} \operatorname{erfc} \left[\frac{x-\xi+rt}{2\sqrt{\alpha t}} \right] d\xi \right. \\
& - \int_{x+rt}^{\infty} \frac{d}{d\xi} \left\{ F(\xi) \left[1 + \exp \left(\frac{-x\xi}{\alpha t} \right) \right] \right\} \operatorname{erfc} \left[\frac{\xi-x-rt}{2\sqrt{\alpha t}} \right] d\xi \\
& \left. + \left(\frac{r}{\alpha} + 2p \right) \int_0^{\infty} F(\xi) \exp \left[\frac{r\xi}{\alpha} + p(x+\xi+rt+p\alpha t) \right] \operatorname{erfc} \left[\frac{x+\xi+rt+2p\alpha t}{2\sqrt{\alpha t}} \right] d\xi \right\}
\end{aligned}$$

$$\begin{aligned}
\frac{\partial T_{i3}}{\partial x} = & F'(x+rt) \left\{ 1 - \exp \left[-\frac{x}{\alpha} \left(\frac{x}{t} + r \right) \right] \right\} \\
& + p \left\{ 2F(x+rt) \exp \left[\frac{-x}{\alpha} \left(\frac{x}{t} + r \right) \right] - F(0) \operatorname{erfc} \left[\frac{x+rt}{2\sqrt{\alpha t}} \right] \right\} \\
& - \int_0^{x+rt} \frac{d}{d\xi} \left\{ pF(\xi) \exp \left(\frac{-x\xi}{\alpha t} \right) + \frac{1}{2} F'(\xi) \left[1 - \exp \left(\frac{-x\xi}{\alpha t} \right) \right] \right\} \operatorname{erfc} \left[\frac{x-\xi+rt}{2\sqrt{\alpha t}} \right] d\xi \\
& + \int_{x+rt}^{\infty} \frac{d}{d\xi} \left\{ pF(\xi) \exp \left(\frac{-x\xi}{\alpha t} \right) + \frac{1}{2} F'(\xi) \left[1 - \exp \left(\frac{-x\xi}{\alpha t} \right) \right] \right\} \operatorname{erfc} \left[\frac{\xi-x-rt}{2\sqrt{\alpha t}} \right] d\xi \\
& - \frac{p}{2} \left[\frac{r}{\alpha} + 2p \right] \int_0^{\infty} F(\xi) \exp \left[\frac{r\xi}{\alpha} + p(x+\xi+rt+p\alpha t) \right] \operatorname{erfc} \left[\frac{rt+2p\alpha t+x+\xi}{2\sqrt{\alpha t}} \right] d\xi
\end{aligned}$$

$$\begin{aligned}
T_{i3}(0, t) = & 2F(rt) - F(0) \operatorname{erfc} \left[\frac{r}{2} \sqrt{\frac{t}{\alpha}} \right] - \int_0^{rt} F'(\xi) \operatorname{erfc} \left[\frac{rt-\xi}{2\sqrt{\alpha t}} \right] d\xi \\
& + \int_{rt}^{\infty} F'(\xi) \operatorname{erfc} \left[\frac{rt-\xi}{2\sqrt{\alpha t}} \right] d\xi - \left(\frac{r}{2\alpha} + p \right) \\
& \times \int_0^{\infty} F(\xi) \exp \left[\frac{r\xi}{\alpha} + p(\xi+rt+p\alpha t) \right] \operatorname{erfc} \left[\frac{\xi+rt+2p\alpha t}{2\sqrt{\alpha t}} \right] d\xi
\end{aligned}$$

$$\begin{aligned}
\frac{\partial T_{i3}}{\partial x}(0, t) = & p \left\{ 2F(rt) - F(0) \operatorname{erfc} \left[\frac{r}{2} \sqrt{\frac{t}{\alpha}} \right] \right\} \\
& - \int_0^{rt} F'(\xi) \operatorname{erfc} \left[\frac{rt-\xi}{2\sqrt{\alpha t}} \right] d\xi + \int_{rt}^{\infty} F'(\xi) \operatorname{erfc} \left[\frac{\xi-rt}{2\sqrt{\alpha t}} \right] d\xi \\
& - \left[\left(\frac{r}{2\alpha} \right) + p \right] \int_0^{\infty} F(\xi) \exp \left[\frac{r\xi}{\alpha} + p(\xi+rt+p\alpha t) \right] \\
& \times \operatorname{erfc} \left[\frac{\xi+rt+2p\alpha t}{2\sqrt{\alpha t}} \right] d\xi
\end{aligned}$$

$$T_{i3}(x, 0) = F(x)$$

$$\frac{\partial T_{i3}}{\partial x}(x, 0) = F'(x)$$

Table 2 (contd)

T_{g3} (heat generation or radiation penetration)	
$T_{g3} = \frac{1}{2} \int_0^t \int_0^\infty G(\xi, t - \lambda) \left\{ \sqrt{\frac{\alpha}{\pi\lambda}} \exp\left[-\frac{(x - \xi + r\lambda)^2}{4\alpha\lambda}\right] \left[1 + \exp\left(\frac{-x\xi}{\alpha\lambda}\right)\right] \right. \\ \left. - (r + 2\alpha p) \exp\left[\frac{r\xi}{\alpha} + p(x + \xi + r\lambda + \alpha p\lambda)\right] \operatorname{erfc}\left[\frac{x + \xi + r\lambda + 2p\alpha\lambda}{2\sqrt{\alpha\lambda}}\right] \right\} d\xi d\lambda$ $\frac{\partial T_{g3}}{\partial x} = -\frac{1}{2} \int_0^t \int_0^\infty G(\xi, t - \lambda) \left\{ \frac{1}{2\sqrt{\pi\alpha\lambda^3}} \exp\left[-\frac{(x - \xi + r\lambda)^2}{4\alpha\lambda}\right] \left[(x - \xi + r\lambda) \right. \right. \\ \left. \left. + (x + \xi - r\lambda - 4\alpha p\lambda) \exp\left(\frac{-x\xi}{\alpha\lambda}\right) \right] - p(r + 2\alpha p) \exp\left[\frac{r\xi}{\alpha} + p(x + \xi + r\lambda + \alpha p\lambda)\right] \right. \\ \left. \operatorname{erfc}\left[\frac{x + \xi + r\lambda + 2p\alpha\lambda}{2\sqrt{\alpha\lambda}}\right] \right\} d\xi d\lambda$	
$T_{g3}(0, t) = \int_0^t \int_0^\infty G(\xi, t - \lambda) \left\{ \sqrt{\frac{\alpha}{\pi\lambda}} \exp\left[-\frac{(r\lambda - \xi)^2}{4\alpha\lambda}\right] \right. \\ \left. - \left(\frac{r}{2} + \alpha p\right) \exp\left[\frac{r\xi}{\alpha} + p(\xi + r\lambda + \alpha p\lambda)\right] \operatorname{erfc}\left[\frac{\xi + r\lambda + 2p\alpha\lambda}{2\sqrt{\alpha\lambda}}\right] \right\} d\xi d\lambda$ $\frac{\partial T_{g3}}{\partial x}(0, t) = -p \int_0^t \int_0^\infty G(\xi, t - \lambda) \left\{ \sqrt{\frac{\alpha}{\pi\lambda}} \exp\left[\frac{(r\lambda - \xi)^2}{4\alpha\lambda}\right] \right. \\ \left. - \left(\frac{r}{2} + \alpha p\right) \exp\left[\frac{r\xi}{\alpha} + p(\xi + r\lambda + \alpha p\lambda)\right] \operatorname{erfc}\left[\frac{\xi + r\lambda + 2p\alpha\lambda}{2\sqrt{\alpha\lambda}}\right] \right\} d\xi d\lambda$	$T_{g3}(x, 0) = 0$ $\frac{\partial T_{g3}}{\partial x}(x, 0) = 0$
Expansion of the condensed functions	
$\frac{\partial}{\partial t} \int_0^t P(t - \lambda) Q(x, \lambda) d\lambda = P(0) Q(x, t) + \int_0^t \frac{\partial P}{\partial \lambda}(\lambda) Q(x, t - \lambda) d\lambda$	
Definition of the complementary error function	
$\operatorname{erfc}(\lambda) = \frac{2}{\sqrt{\pi}} \int_\lambda^\infty \exp(-\xi^2) d\xi$	

Table 2 also contains the solutions to the differential equation with the various types of boundary conditions. The solutions were obtained by taking Laplace transforms of the equation with respect to time, then solving the resulting ordinary differential equations in a straightforward manner, using the boundary conditions to evaluate the integration constants, and finally transforming back into the (x, t) plane.

The solutions are broken down into contributions due to surface boundary condition, initial temperature distribution, and heat generation in the solid. The solutions shown in the table are for constant r and p . All the integrals are proper wherever they converge and lead to the correct limits as either x or t approach zero. The three T_i functions are given both in compact form and also in an expanded form that is useful for numerical computation

in that it leads to the proper limits at x and t approaching zero. The first form of T_{s1} cannot be evaluated by the equation shown for $r = 0$. However, an alternate form is shown in Table 2 for $r = 0$. Thus the latter equation should be used in a numerical evaluation of T_{s1} for $r = 0$. With this exception the equations shown may be used for all r .

3. Extension to Time-Varying Surface Motion and Convection

The most obvious way to extend these equations to the theoretically important cases of time-varying r and p is to choose a set of time intervals which are sufficiently small so that these two parameters may be considered constant over the interval (Fig. 6). The equations would then be evaluated over the first such interval, the resulting solution used as an initial condition for the second interval and a solution obtained for that interval, and so forth for each small interval. This is a laborious method, however, particularly in view of the fact that one is generally only interested in following the surface temperature and its first derivative with respect to x . The method suggested above would involve calculation of temperatures over the whole range of x .

Fortunately there is a less laborious approach. Suppose r and p can be considered to be constant over sufficiently small intervals in time. That is

$$\begin{aligned} r &= r_i & t^{(i)} \leq t < t^{(i+1)} \\ p &= p_i & t^{(i)} \leq t < t^{(i+1)} \end{aligned}$$

where the subscript on r and p and superscript on t designate the particular interval. Then for a given time interval it is possible to make a change of variables in the equations such that x , r and p appear explicitly only in the function F , G and H , and if Case 3 boundary conditions are used, in a constant in that equation.

The change of variables is:

$$\begin{aligned} \xi_n &= \frac{r_n x}{\alpha} \\ \tau_n &= \frac{r_n^2 t}{\alpha} \end{aligned}$$

which leads to the equation

$$\frac{\partial^2 T_n^*}{\partial \xi_n^2} + \frac{\partial T_n^*}{\partial \xi_n} = \frac{\partial T_n^*}{\partial \tau_n} - G_n^*(\xi_n, \tau_n)$$

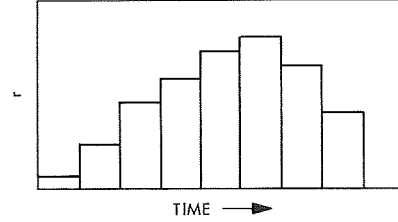


Fig. 6. Representation of variable burning rate with time

with boundary conditions

$$T_n^*(\xi_n, 0) = F_n^*(\xi_n)$$

$$T_n^*(\infty, \tau_n) = 0$$

and

Case 1

$$\frac{\partial T_n^*}{\partial \xi_n}(0, \tau_n) = -H_n^{(1)}(\tau_n)$$

Case 2

$$T_n^*(0, \tau_n) = H_n^{(2)}(\tau_n)$$

Case 3

$$T_n^*(0, \tau_n) - K_n \frac{\partial T_n^*}{\partial \xi_n}(0, \tau_n) = H_n^{(3)}(\tau_n)$$

where the functions are defined as follows:

$$G_n^*(\xi_n, \tau_n) = \left(\frac{\alpha}{r_n}\right)^2 G\left(\frac{\alpha \xi_n}{r_n}, \frac{\alpha \tau_n}{r_n^2}\right)$$

$$F_n^*(\xi_n) = F\left(\frac{\alpha \xi_n}{r_n}\right)$$

$$H_n^{(1)}(\tau_n) = \left(\frac{\alpha}{r_n}\right) H\left(\frac{\alpha \tau_n}{r_n^2}\right)$$

$$H_n^{(2)}(\tau_n) = H\left(\frac{\alpha \tau_n}{r_n^2}\right)$$

$$H_n^{(3)}(\tau_n) = H\left(\frac{\alpha \tau_n}{r_n^2}\right)$$

$$K_n = \frac{r_n}{p_n \alpha}$$

The method of solution may now be developed by mathematical induction. Suppose that for a set of functions, $F_n^*, G_n^*, H_n^{(i)}$ ($i = 1, 2$, or 3) and a constant K_n , a solution is found which is valid in the interval

$$0 \leq \tau_n \leq \frac{r_n^2 t^{(n)}}{\alpha}$$

and which is designated $T_n^*(\xi_n, \tau_n)$. Suppose further that when this solution is transformed back into the x - t plane, it yields the correct solution to the original differential equation and boundary conditions at $t = t^{(n)}$. That is, $T_n(x, t^{(n)})$ is the correct solution at $t^{(n)}$. We will then show how the correct solution at $t^{(n+1)}$ may be obtained.

If the problem is being solved for a Case 3 boundary condition, we begin by changing that condition to

$$T_n^*(0, \tau_n) - K_{n+1} \frac{\partial T_n^*}{\partial \xi_n}(0, \tau_n) = H_n^{(4)}(\tau_n)$$

where

$$H_n^{(4)}(\tau_n) = H_n^{(3)}(\tau_n) + (K_n - K_{n+1}) \frac{\partial T_n^*}{\partial \xi_n}(0, \tau_n)$$

Since presumably the solution has already been obtained for the n th interval, $\partial T_n^* / \partial \xi$ is known. If either of the other two types of boundary conditions are considered, this step is omitted.

Now if a similar transformation is made for $(n+1)$ instead of n , we note first of all that the resulting equations correspond in the $\xi_{n+1} - \tau_{n+1}$ plane to the correct original equations in the x - t plane in the interval $t^{(n)} \leq t < t^{(n+1)}$. We would like the solution with subscripts $(n+1)$ to match the solution T_n at $t^{(n)}$ (Fig. 7). That is

$$T_{n+1}(x, t^{(n)}) = T_n(x, t^{(n)})$$

from which it follows that

$$T_{n+1}\left(\frac{\alpha \xi_{n+1}}{r_{n+1}}, \frac{\alpha \tau_{n+1}}{r_{n+1}^2}\right) = T_n\left(\frac{\alpha \xi_n}{r_n}, \frac{\alpha \tau_n}{r_n^2}\right)$$

$$T_{n+1}^*(\xi_{n+1}, \tau_{n+1}) = T_n^*(\xi_n, \tau_n)$$

But the form of the transformed equations for $(n+1)$ is identical with that for (n) , so that the above condition is satisfied if

$$F_{n+1}^*(\xi_{n+1}) = F_n^*(\xi_n)$$

$$G_{n+1}^*(\xi_{n+1}, \tau_{n+1}) = G_n^*(\xi_n, \tau_n)$$

$$H_{n+1}^{(i)}(\tau_{n+1}) = H_n^{(j)}(\tau_n) \quad (i = 1, 2, 3; j = 1, 2, 4)$$

in the interval $0 \leq t < t^{(n)}$. These equations can be shown to reduce to the following requirements:

$$F_{n+1}^*(\xi_{n+1}) = F\left(\frac{\alpha \xi_{n+1}}{r_{n+1}}\right)$$

$$G_{n+1}^*(\xi_{n+1}, \tau_{n+1}) = \left(\frac{\alpha}{r_i}\right)^2 G\left(\frac{\alpha \xi_{n+1}}{r_{n+1}}, \frac{\alpha \tau_{n+1}}{r_{n+1}^2}\right), \quad \frac{r_{n+1}^2 t^{(i)}}{\alpha} \leq \tau_{n+1} < \frac{r_{n+1}^2 t^{(i+1)}}{\alpha} \quad (i = 1, 2, \dots, n)$$

$$H_{n+1}^{(1)}(\tau_{n+1}) = \left(\frac{\alpha}{r_i}\right) H\left(\frac{\alpha \tau_{n+1}}{r_{n+1}^2}\right), \quad \frac{r_{n+1}^2 t^{(i)}}{\alpha} \leq \tau_{n+1} < \frac{r_{n+1}^2 t^{(i+1)}}{\alpha} \quad (i = 1, 2, \dots, n)$$

$$H_{n+1}^{(2)}(\tau_{n+1}) = H\left(\frac{\alpha \tau_{n+1}}{r_{n+1}^2}\right)$$

$$H_{n+1}^{(3)}(\tau_{n+1}) = \begin{cases} H\left(\frac{\alpha \tau_{n+1}}{r_{n+1}^2}\right) + (K_n - K_{n+1}) \left(\frac{\alpha}{r_n}\right) \frac{\partial T_n}{\partial x}\left(0, \frac{\alpha \tau_{n+1}}{r_{n+1}^2}\right), & 0 \leq \tau_{n+1} \leq \frac{r_{n+1}^2 t^{(n)}}{\alpha} \\ H\left(\frac{\alpha \tau_{n+1}}{r_{n+1}^2}\right), & \frac{r_{n+1}^2 t^{(n)}}{\alpha} \leq \tau_{n+1} < \frac{r_{n+1}^2 t^{(n+1)}}{\alpha} \end{cases}$$

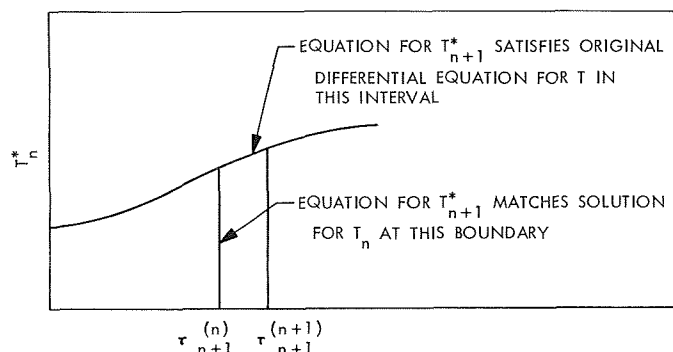


Fig. 7. Method of matching solutions

It is also clear that this entire method applies to the first interval, which completes the mathematical induction reasoning. The accuracy of the solutions may be increased by decreasing the interval size.

To sum up, then, to obtain the solution at $t^{(n+1)}$, one uses the appropriate solution for the transformed equation (which is available from Table 2) evaluated at

$$\tau_{n+1} = \frac{r_{n+1}^2 t^{(n+1)}}{\alpha}$$

with the functions F^* , G^* , and $H^{(i)}$ evaluated as indicated in the equations above. This solution is then transformed back to the $x-t$ plane. The advantage of this method is that if a temperature (or its derivative) is only required at the surface, no information about the temperature profile at other locations, with the exception of the initial temperature distribution, need be retained in the calculation as it proceeds in increasing time.

4. Application to Transient Burning of Solid Propellants

The equations presented in the preceding section apply when r and p are known *a priori* as functions of time. In the description of transient burning of solid propellants this is usually not the case. The rate of motion of the surface, which is equal to the burning rate of the propellant is usually linked in some way with the temperature at the surface, which usually is not known *a priori*. The most common expression for surface regression of a burning solid is $r = A \exp(-B/T_0)$ where T_0 is the surface temperature and A and B are constants. This expression is in the form of the well-known Arrhenius expression for the rate of chemical reactions, and the rationale for its use is that a kinetically controlled chemical reaction is taking place at the surface which produces the decomposition.

The fact that the rate of surface motion is related to temperature at the surface complicates the problem, because neither are known explicitly at a given time. There are undoubtedly many possible methods for performing the necessary calculations, one of which is suggested here. If the above equation is differentiated, one obtains

$$\frac{dr}{dt} = \frac{B}{T_0^2} \left(\frac{\partial T_0}{\partial t} \right) r$$

According to the reasoning of the previous section, T_0 can be thought of and calculated as a function of r and t . Thus the above equation can be thought of as a differential equation for r with t as the independent variable and can be solved by a number of well-known techniques. Thus if r and T_0 are known at some time equal to $t^{(n)}$, the differential equation may be used to obtain both r and T_0 at $T^{(n+1)}$ simultaneously. Through the use of implicit techniques, accuracy may be maintained throughout the course of the calculations.

C. Functionality Determination of Binder Prepolymers by Thin Layer Chromatography,

H. E. Marsh, Jr., and W. M. Foley, Jr.

1. Introduction

For some time, JPL has been interested in the determination of the functionality distribution of solid propellant binder prepolymers. The major thrust has been the adaption of network polymer phenomena (see Subsection XVII-D). Others in the propellant industry have approached the problem by attempting to physically separate prepolymer fractions according to the number of functional groups (hydroxyl or carboxyl) per molecule. The most successful physical separation method has been developed at ESSO by Muenker and Hudson (Ref. 1). They use liquid chromatographic separation on a silica gel column. A quicker, but up to this point nonquantitative, method has been explored and put to practice by Ashcraft (Ref. 2). In this case, thin layer chromatography (TLC) is the technique. Again, silica gel is used, on glass plate backing.

Investigation of TLC has begun at JPL with two purposes in mind. The first purpose is to fulfill part of a goal to characterize and compare a rather large variety of propellant binder prepolymers. The second purpose is to provide a means of checking and supplementing the method of functionality distribution determination based on network polymer phenomena.

2. Instant Thin Layer Chromatography

Current studies have concentrated on using silica gel-impregnated glass fiber sheets in place of the glass plates mentioned above. As the name instant thin layer chromatography (ITLC) implies, the elution process is much more rapid with the fiber sheets. Although solvent matching is different from one plate type to another, once a good system is found for a given plate-prepolymer combination, other differences between plate types can be seen. Insufficient study has been made in this case to reach a positive conclusion, but qualitatively it appears that the resolution between fractions is a little higher on fiber sheets than on glass plates.

Telagen-s Lot 242AM-293/299H (a General Tire & Rubber Co. hydroxyl-terminated polybutadiene prepolymer) was available in two samples, the original preparation which would not gel in a standard test and a reprocessed sample which did gel properly. Each sample was dissolved in chloroform (1 g in 10 ml) and 2-lambda aliquots were placed on an ITLC plate. The chromatogram was developed in a mixture of 10% benzene in hexane, and it showed that the sample which gelled properly was mainly difunctional, while the sample which failed to gel had much more monofunctional material and even a small amount of a nonfunctional fraction. The reprocessed sample gave an R_f for the monofunctional fraction of 0.86 and an R_f of the difunctional fraction of 0.49, while the original sample (the one which failed to gel) gave an R_f of the monofunctional of 0.85 an R_f of the difunctional fraction of 0.44 and a small amount of nonfunctional material with an R_f of 1.0.

Telagen-s Lot 242AM-148-AH, which was one of the best Telagen lots with respect to propellant properties, was evaluated on a silica gel ITLC plate to show three fractions: nonfunctional, monofunctional, and the major fraction difunctional prepolymer. The nonfunctional fraction had an R_f of 1.0, the monofunctional fraction an R_f of 0.76 and 0.77, and the difunctional fraction an R_f of 0.36 and 0.37.

The significant feature of these preliminary analyses is that both the good lot and the reprocessed portion of the bad lot have a high fraction of difunctional component. This correlates with the fundamental concepts of network polymer theory.

References

1. Muenker, A. H., and Hudson, B. E., *Functionality Determination of Binder Prepolymers*, Esso Report GR-8-FBP-68, p. 27,

Final Report for the period Oct. 1, 1966 to Sept. 30, 1968 (AFRPL-TR-68-237, Contract F04611-67-C-0012), Lynden, N.J.

2. Ashcraft, A. C., et al., *Saturated Hydrocarbon Polymeric Binder for Advanced Solid Propellants*, Triennial Report for the period Oct. 27, 1965 to Sept. 30, 1968, Union Carbide Corp., under JPL Contract 951210, Bound Brook, N.J., Oct. 31, 1968.

D. Gelation Studies, H. E. Marsh, Jr., and J. F. Wolfe

1. Introduction

This work is a continuation of a series of investigations of network polymer formation reported earlier (SPS 37-42, Vol. IV, p. 106; SPS 37-43, Vol. IV, p. 163; SPS 37-45, Vol. IV, p. 77; SPS 37-47, Vol. III, p. 69; SPS 37-48, Vol. III, p. 95; SPS 37-49, Vol. III, p. 180; and Ref. 1). Stimulated by problems encountered in efforts to develop better binders for solid propellants, this work has led to some fundamental extensions of Flory's network theory (Ref. 2; and SPS 37-48, Vol. III, p. 95) which may be of use to other fields concerned with branched, step-wise polymerizations. The chief purpose of the present work is the development of a method of determining the functionality distribution of propellant binder prepolymers. The current effort deals with the simplest of gelling systems in order to lay the groundwork for testing the theoretical extension mentioned above. A parallel effort using different experimental techniques is going on under a JPL-managed NASA contract (NAS7-689) by J. Heller and co-workers at Stanford Research Institute (Refs. 3, 4, and 5).

2. Theoretical Models

The theoretical characterization of network polymers generated as the reaction products of polyfunctional molecules discussed previously (Ref. 1; and SPS 37-48, Vol. III, p. 95) is reviewed here briefly. The simple system under consideration is polyesterification of tri- and bifunctional acids with bifunctional alcohols carried out in xylene solution.

Flory's theory (Ref. 2) states that a reaction formulation will not gel until a critical value of α , the branching probability, has been reached. α can be defined explicitly in terms of the extent of reaction and the stoichiometric ratios of the starting materials. Incipient gelation is said to occur when this critical value of α has been reached, which is marked by the first appearance of an insoluble gel. For the simplest gelling system (taken to complete reaction), theory states that incipient gelation occurs at $R = 1 + \rho$, where R is the ratio of all hydroxyls to all carboxyls and ρ is the ratio of carboxyls on branch units to the total carboxyls. If $R > 1 + \rho$ the system will remain liquid at complete reaction. At values of R less than $1 + \rho$,

infinite networks are possible, and in xylene solution this is observed as a plasticized gel.

3. Model Experiments

Polyesterifications of the present study are carried out in xylene solvent at about 50% concentration. The water of reaction is removed by azeotropic distillation; xylene is returned to the reaction vessel by means of a decanting receiver. Methods of incipient gelation detection will be discussed below. By contrast, in the earliest work on this program reactions were carried out in bulk, and water removal depended on diffusion through pinholes into a slow flow of dry nitrogen. Incipient gelation, in this case, was identified by non-flowability at 170°C. Still another set of conditions is being investigated by Heller, et al., with considerable success (Refs. 3, 4, and 5). In this case, bulk reactions are stirred very rapidly with a stream of dry nitrogen sweeping the surface. Incipient gelation is detected as a marked rise in viscosity, which is monitored continuously. The time to reach complete reaction was reduced from days to hours by both of the newer methods—stirring and azeotropic distillation.

Heller and coworkers demonstrated that titration from non-gelling compositions to the point of incipient gelation could be employed. Titration is carried out by small incremental additions alternated by stirring at temperature for sufficient time for the reaction to be complete, indicated by a leveling of viscosity. Incremental titration is being applied also to the solution studies, reported here. In this case, completion of reaction is indicated by discontinuation of water distillation.

The experimental studies of solution polyesterification are conducted as follows: Reactions are carried out in a 1-liter resin flask with mechanical stirrer, heated with an oil bath to 160°C. A flow of dry nitrogen before reflux reduces the amount of discoloration due to oxidation by the air and is also used during later stages of reaction to aid distillation. The byproduct water is removed azeotropically into a distillation receiver with xylene overflowing into the reaction flask. Analysis of water and xylene distillate showed no loss of reactants during distillation. Near the end of reaction, when water distillation appears to cease, all of the two-phase azeotrope is removed from the apparatus, and a compensating quantity (equal to the xylene portion) of fresh, dry xylene is added to the kettle; the process is continued for 4 to 6 h. *p*-Toluene sulfonic acid is used at 0.2–0.5 equivalent percent as catalyst. The system of 1,3,5-pentanetricarboxylic acid, sebacic acid, and ethylene glycol gave results which agree well with

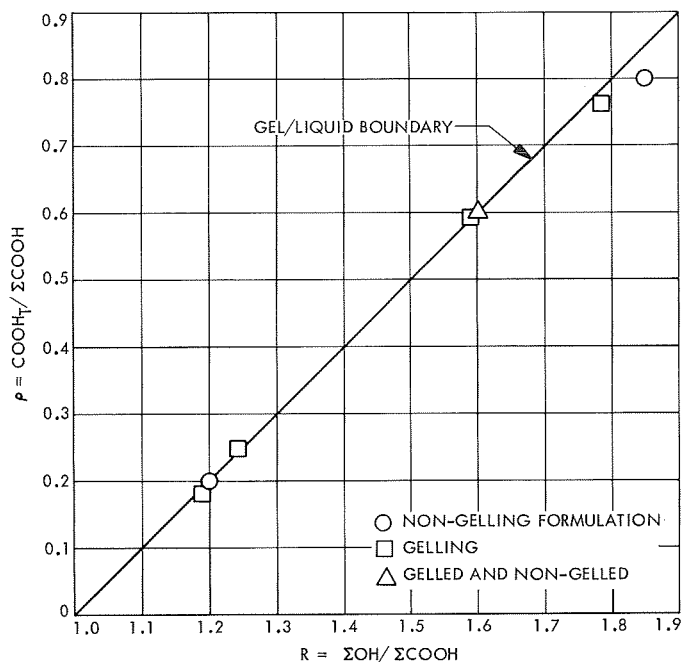


Fig. 8. Incipient gelation boundary for the system: 1,3,5-pentanetricarboxylic acid, sebacic acid, and ethylene glycol

theory for the range of ρ studied. Infrared analysis indicated high extents of reaction.

These results are shown in Fig. 8. The square points indicate observation of gelation, while the circles represent compositions which did not gel.

4. Conclusion

End-point determination with the 1,3,5-pentanetricarboxylic acid/sebacic acid/ethylene glycol system appears to be unambiguous. Compositions at the incipient gelation boundary remained liquid for several hours of reaction. Then gelation occurred rapidly, visibly noticeable as a large mass of elastic material. Reference mixtures, theoretically liquid, but very close to gelling, remained liquid for hours after water distillation stopped. Visually observed gels are confirmed by soxhlet extractions.

The results are less clear cut with a similar system, in which the ethylene glycol is replaced with 1,10-decanediol. It appears necessary to carry reactions beyond the theoretical incipient gelation boundary to get visually observable gels. Under question are several of the assumptions: homogeneity of reaction mixture, equal reactivity regardless of molecular size, the absence of closed loops. In particular, evidence of the existence (or nonexistence) of microgel particles is being looked for.

References

1. Marsh, H. E., Jr., and Hutchison, J. J., "Functionality Determination Through Flory Network Theory," Chemical Propulsion Information Agency Publication 187, p. 55, Johns Hopkins University, Silver Spring, Md., March 1969 (Confidential).
2. Flory, P. J., *J. Am. Chem. Soc.*, Vol. 63, p. 3083, 1941.
3. Hodgkin, J. H., Martinelli, F. J., and Heller, J., *Project No. PRU 7471*, Quarterly Report No. 1, Stanford Research Institute, Menlo Park, Calif., 1968.
4. Hodgkin, J. H., Martinelli, F. J., and Heller, J., *Project No. PRU 7471*, Quarterly Report No. 2, Stanford Research Institute, Menlo Park, Calif., 1969.
5. Hodgkin, J. H., Martinelli, F. J., and Heller, J., *Project No. PRU 7471*, Quarterly Report No. 3, Stanford Research Institute, Menlo Park, Calif., 1969.

E. Electronic Conductivity of Solid Propellant

Exhaust Products, O. K. Heiney

1. Introduction

The central problem in developing a solid-rocket-driven magnetohydrodynamic generator is the optimization of the electrical conductivity of the exhaust gases. If possible, it is greatly advantageous to use analytic rather than experimental techniques to determine trends in conductivities as a function of the constituent ingredients of the propellant formulations. This article describes how these conductivities may be determined from simple kinetic theory and experimentally determined collision cross sections. Mole fractions of combustion products and ionized species are determined by shifting equilibrium free energy minimization techniques. Finally, combustion products and their contributive cross sections are presented for a typical case. Terms used are defined in Table 3.

2. Analysis

Given an electron concentration in a gas, a value for the conductivity may be calculated from kinetic theory in the following manner:

From Ref. 1 the conductivity is shown to be simply

$$\sigma = \frac{n_e e^2 \tau}{m_e} \quad (1)$$

from equating the current density and electron drift equations.

τ is the mean free time between electron collisions and is equal to the reciprocal of the mean thermal velocity

Table 3. Definition of terms

C	= rms gas molecular velocity
\bar{c}	= mean gas molecular velocity
e	= charge of electron
K	= Boltzmann constant
m_e	= mass of electron
m_i	= molecular mass of i th species gas
n_e	= number of electrons
n_i	= number of i th species
n_0	= total number of particles
N_0	= Avogadro's Number
P	= gas pressure
P_w	= power output
Q_{ie}	= cross-section of i th species for electron collision
R	= gas constant
T	= gas temperature
U	= gas flow velocity
σ	= gas electrical conductivity
τ	= mean time between electron collision with gas molecules

and the summation of the electron collision cross sections or

$$\tau = \frac{1}{n_0 \bar{c} \sum_i \frac{n_i}{n_0} Q_{ie}} \quad (2)$$

Then from kinetic theory (Ref. 2),

$$PV = \frac{1}{3} n_0 m_i C^2 \quad (3)$$

while equation of state gives

$$PV = \frac{n_0}{N_0} RT = n_0 KT \quad (4)$$

Combining Eqs. (3) and (4) gives

$$C = \left(\frac{3KT}{m_i} \right)^{1/2} \quad (5)$$

However, this C is a root mean square velocity when the required value is for \bar{c} or the mean velocity. Reference 2 gives the relation between these velocities to be

$$\frac{\bar{c}}{C} = \left(\frac{8}{3\pi} \right)^{1/2}$$

which when substituted into Eq. (1) with Eq. (2) results in

$$\sigma = \frac{e^2 (\pi/8)^{1/2}}{(m_e K T)^{1/2} \sum_i \frac{Q_i n_i}{n_e}}$$

or evaluating constants

$$\sigma = \frac{4.55 \times 10^{-12}}{T^{1/2} \sum_i \frac{Q_i n_i}{n_e}} \text{ mho/m}$$

For conductivity determination of any given propellant formulation the values of T , m_i , and n_e may be determined from free-energy computations, while the values of Q_i are determined experimentally and tabulated in Ref. 3. Most properly they should be considered as a function of temperature; however, for the purposes required here, a single value will be adequate. It should be noted that for the potassium used for the electron donor, recent advances in free-energy computer thermochemical programs allow the free-electron concentration to be computed directly rather than through a Saha-type ionization equation. This approach also computes the concentration of species which will capture electrons such as CO_2 , OH , NO_2 , O and reduce the net electron density.

The values used for the collision cross section of the combustion products are given in Table 4.

The propellant formulation which is considered in the figures and discussion below was a double base, HMX,² aluminum, and potassium nitrate system. The oxidizing ingredients are of a class A type for two compelling reasons: (1) the combustion temperature must be fairly high to obtain significant potassium ionization, and (2) perchlorate-type oxidizers must be avoided to eliminate chlorine from the exhaust. The affinity of the chlorine for free electrons is such that they would be virtually all adsorbed.

Fig. 9 illustrates the concentration of the nonionized species in the propellant exhaust as a function of area ratio or distance in nozzle, assuming shifting equilibrium.

²HMX = $\text{C}_4\text{H}_8\text{N}_8\text{O}_8$.

Table 4. Species collision cross section

Specie	Cross section, $\text{m}^2 \times 10^{-20}$
CO	8
CO ₂	14
H	58
H ₂	13
H ₂ O	60
N	18
N ₂	11
NO	6
O	10
O ₂	4.6
OH	51
K ⁺	5,000
AlO	10
Al ₂ O ₃	10
K	520
KO	10
KOH	10

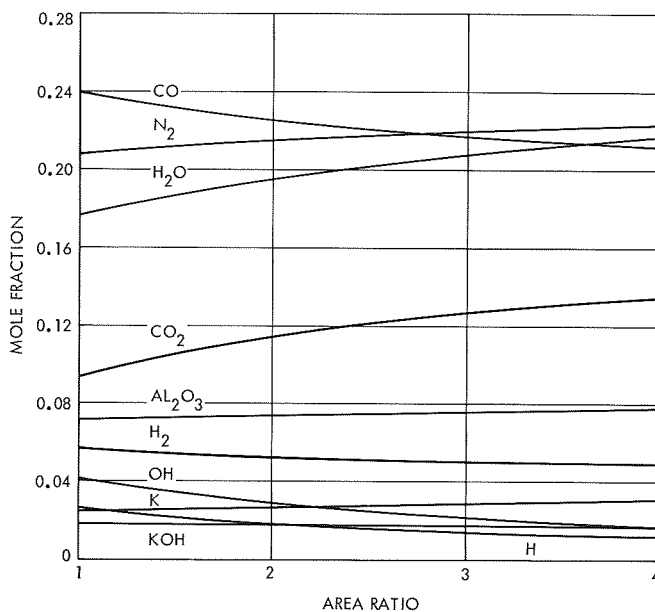


Fig. 9. Mole fraction of combustion products as function of area ratio

Figure 10 is the same relation for the ionized species, with the curve labeled E being that for the free electrons. The fraction of potassium ionized decreases as the area ratio increases and is largely due to the progressively lowering temperature. An interesting feature is that the greatest capture of the electrons is by CO_2 rather than OH .

Prior work in this field often ignores all species other than OH . It is seen that this leads to significant error, particularly at higher temperatures.

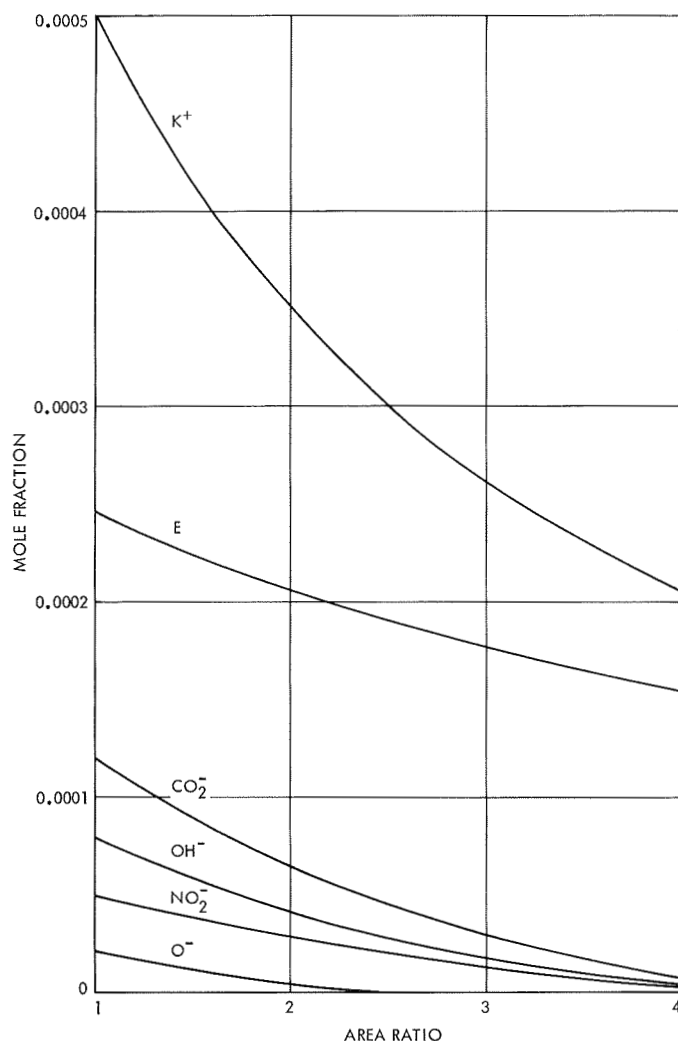


Fig. 10. Mole fraction of ionized species as function of area ratio

The percent contribution to total electron cross section is shown on Fig. 11. It is apparent that water vapor and potassium are the major impeding elements.

The power output of a solid propellant MHD generator is not only a function of the conductivity of the gases σ but also of the square of their velocity. Thus

$$P_w \propto \sigma U^2$$

and the term σU^2 is the one to be optimized rather than σ . The conductivity and σU^2 results of the formulation given above, at various area ratios, is as shown in Table 5.

These σU^2 values of greater than 1.5×10^8 mho-m/s compare very favorably to existing liquid fuel MHD

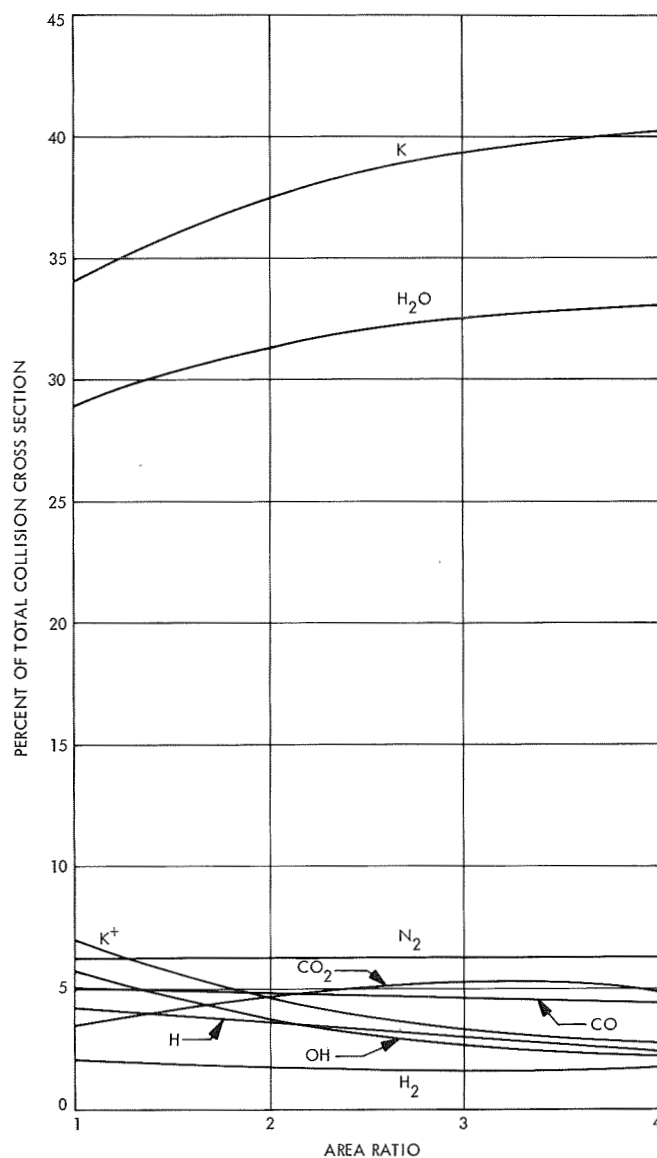


Fig. 11. Percent of species contribution to total collision cross section

Table 5. Results of conductivity computations

Area ratio	Mach No.	Temperature, °K	Electron mole fraction $\times 10^{-5}$	Conductivity, mho/m $\times 10^5$	σU^2 , mho — m/s ²
1.00	1.0	3600	24	48.7	0.49
1.17	1.4	3430	23	47.3	0.92
1.55	1.8	3275	21	43.7	1.22
1.73	1.9	3225	21	43.8	1.35
2.09	2.0	3140	20	42.0	1.47
2.44	2.2	3080	19	40.1	1.54
3.27	2.4	2975	17	36.2	1.60
4.05	2.5	2900	16	34.3	1.65
10.46	3.1	2575	9	20.0	1.28

generators that demonstrate values in the 10^7 range. Solids of similar formulation experimentally have shown σU^2 values (Ref. 4) in this range, while the demonstrated liquid driven MHD generator with the 10^7 value was taken from Ref. 5.

3. Conclusions

Realistic analytic optimization of solid propellant exhaust gas conductivity may be simply performed by developed computing techniques. The resultant conductivity values determined are quite promising and show potential for enhanced performance of MHD power generators using a solid propellant rather than a liquid propellant base material.

References

1. Drummond, J. E., "Plasma Physics," McGraw-Hill Book Co., Inc., New York, 1961.
2. Loeb, L. B., "The Kinetic Theory of Gases," Dover Publications, Inc., New York, 1961.
3. Bincer, H., "Electron Concentration and Electrical Conductivity of Seeded Coal Combustion Plasmas," *MHD Electrical Power Generation*, Vol. 1, European Nuclear Energy Association, 1964.
4. Denzel, D. L., et al., "Experimental Study of Diagonal Conducting Wall Generators Using Solid Propellants," *AIAA J.*, Vol. 6, No. 9, 1968.
5. Womack, G. L., et al., "Experimental Studies in Open-Cycle MHD Generation," *MHD Electrical Power Generation*, Vol. 1, European Nuclear Energy Association, 1964.

XVIII. Research and Advanced Concepts

PROPULSION DIVISION

A. Closed-Loop Operation of a Hollow-Cathode

Ion Thruster, E. V. Pawlik and T. D. Masek

1. Introduction

Electron-bombardment ion thrusters that employ a hollow cathode as the primary electron source have been under investigation (SPS 37-57, Vol. III, pp. 180-181, and SPS 37-59, Vol. III, pp. 205-211). Ion thruster performance and lifetime are both improved when using this cathode type, making it desirable in primary propulsion system studies. The control loops necessary for controlling output power and propellant utilization (percent of mercury propellant ionized and accelerated as the ion beam) for thrusters using such cathodes have now been implemented using laboratory power supplies and controllers. Controlled thruster operation has been examined over a 2.7:1 range of output power. Such a range of output power throttling is useful for power matching with a solar array with a varying solar constant.

2. Apparatus

A nominal 20-cm-diam ion thruster, similar to the unit described in SPS 37-59, Vol. III, was used. It differed only

in the ion accelerating grid structure and cathode pole piece. The screen grid was 0.056 cm thick and contained 0.482-cm-diam holes providing a total open area of 75%. A 0.100-cm spacing was maintained between the screen and accelerator grid. The accelerator grid contained 0.356-cm-diam holes. The cathode pole piece construction was modified from that of SPS 37-59, Vol. III, by decreasing the cathode to baffle distance and reducing the pole piece internal volume as suggested in Ref. 1. These modifications improved the thruster operation slightly (decreased ion chamber losses of about 25%), increased the perveance of the ion optical system, and reduced the noise content of the arc current.

The power supplies and control loops used to operate the thruster are shown schematically in Fig. 1. The ion beam-main vaporizer loop was the same as described in Ref. 2. The arc voltage was used to control the cathode flow vaporizer as suggested in the thruster control concepts presented in Ref. 3. A major difference of the present study was a fixed arc voltage set point at a 35-V discharge. A variable current adjustment loop (for power matching) for the arc current was also implemented in the control system. The current setting was adjusted by a function

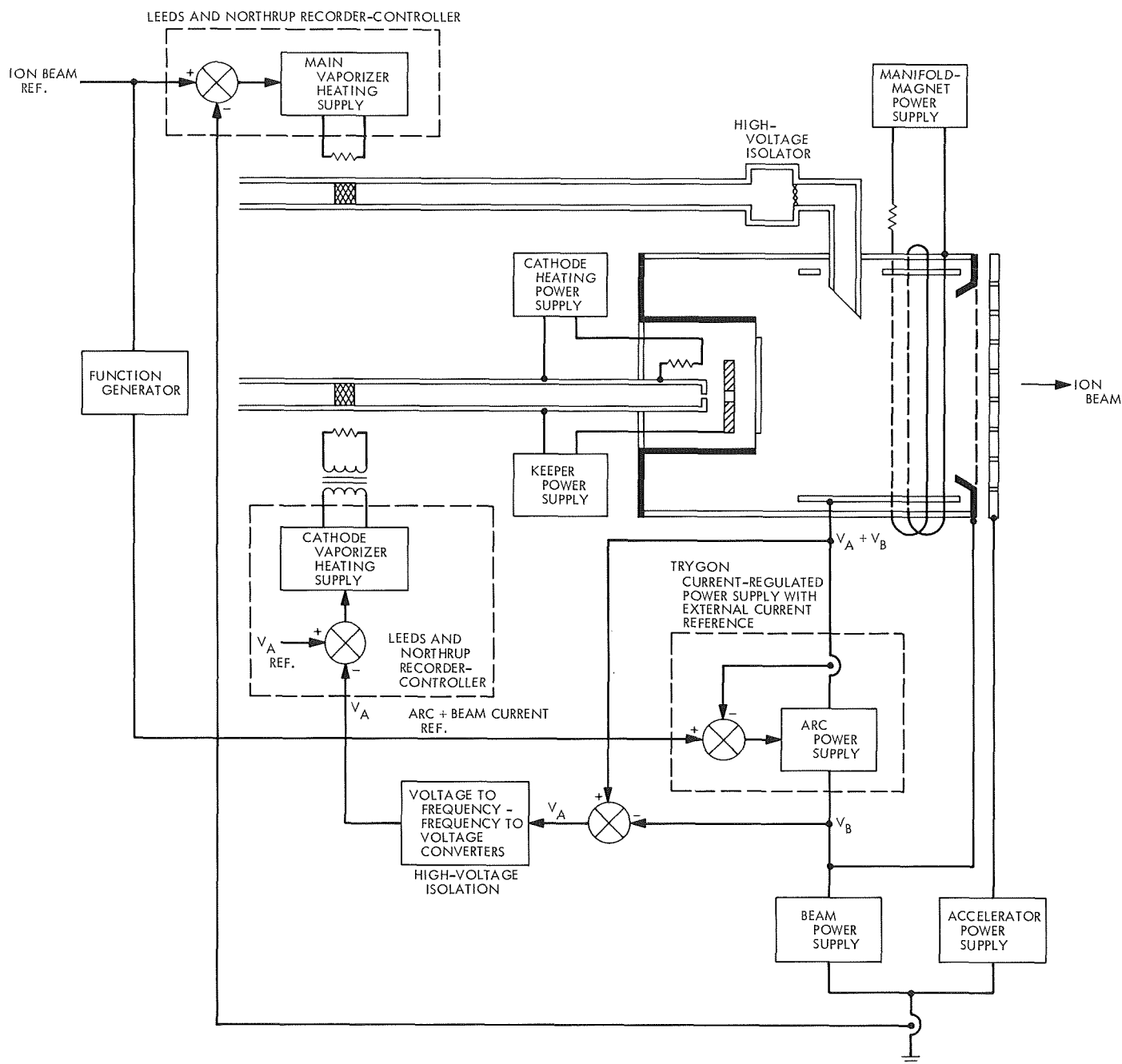


Fig. 1. Schematic of controls for an electron-bombardment ion thruster

generator relating the arc current to the beam current reference. The function generator could be programmed for any of several modes of operation. Two types, operated at constant propellant utilization and at constant arc chamber losses per beam ion, were demonstrated.

Both thruster vaporizers were controlled by Leeds and Northrup recorder-controllers. A Trygon power supply with an external voltage input to set the level of the regu-

lated current was used for the arc power supply. The function generator converted the mechanical set point of the beam-loop Leeds and Northrup recorder-controller into a voltage signal to the arc supply. The second Leeds and Northrup controller was used to maintain a constant arc voltage under varying thruster operating conditions.

The levels of magnet and keeper current were adjusted to provide low noise levels on the arc current. The levels

of these and other power supplies used are listed in Table 1. During all operations the arc voltage reference was maintained at a constant setting of 35 V.

3. Test Results

The control system was operated over a range of output power for two programmed settings of the function generator. These settings correspond to two modes of operation: (1) constant propellant utilization, and (2) constant arc chamber losses per beam ion. Data obtained during this operation are presented in Table 2 and Fig. 2. Each data point was held for a minimum of 1.5 h during which time mercury flow measurements were obtained. The vaporizer temperatures remained constant over this period indicating no drift of the setting. No problems were encountered in either throttling of the output power or maintaining an output power at a constant level. Small errors in the constant propellant utilization or arc chamber loss settings occurred but these could be corrected by further adjustments of the function generator. Both

throttling methods provided about the same penalty in total efficiency.

The response of the system to a sudden step change increase in beam current reference set point is shown in Fig. 3. The arc current increases sharply in response to the change in beam reference. The arc supply current is then regulated about an increased value of arc plus beam current. The arc current, therefore, exhibits a transient fluctuation as the beam current changes. The ion beam current initially increases with the increase in arc current and voltage, decreases as the arc voltage decreases, and finally increases with increasing mercury vapor flow. Steady-state conditions are reached in about 2 min. Some reduction in the transient interaction between control loops should be obtainable if the beam current is subtracted from the arc current control loop. This could not be conveniently accomplished with the present arrangement of laboratory power supplies.

One problem was encountered during the investigation of the control loops. At low values of propellant utilization the ion beam would decrease with increasing mercury vapor flow. Since in this region the beam to flow response curve changed its slope, the thrust could not be controlled. This problem has also occurred for oxide-cathode thruster operation (Ref. 2). The result of disconnecting the function generator and providing an independent reference to the arc and beam control loop is shown in Fig. 4. The arc current was decreased in three successive step decreases to the region where the control loops would not function,

Table 1. Ion thruster control system electric parameters

Parameter	Current, A	Voltage, V
Accelerator	0.0038–0.0062	1000
Beam	0.409–1.110	2000
Magnet–manifold	0.78	17.2
Cathode heat	1.0	2.4
Cathode keeper	0.5	7.0

Table 2. Hollow-cathode ion thruster performance data

Mercury propellant flow rate, equiv amperes	Arc voltage, V	Beam Current, A	Beam Power, W	Power loss, W					Efficiency, %			Operational mode
				Arc chamber	Accelerator	Cathode	Magnet–manifold	Main vaporizer	Power	Propellant utilization	Total	
0.463	35.68	0.409	818	100	18.5	6.4	14	8.9	84.7	88.3	74.8	Constant propellant utilization ^a
0.560	35.42	0.495	990	112	18.3	6.4	14	10.2	86.0	88.4	76.0	
0.843	35.52	0.756	1512	177	14.9	6.4	14	11.4	87.1	89.7	78.2	
1.127	35.51	1.003	2006	238	16.6	6.4	14	12.5	87.7	89.5	78.4	
0.667	36.32	0.610	1220	166	14.4	6.4	14	8.4	85.3	91.5	78.0	Constant arc losses per beam ion ^b
0.865	35.23	0.803	1606	221	11.1	6.4	14	13.0	85.8	92.3	79.3	
1.092	35.08	1.006	2012	273	11.4	6.4	14	12.0	86.5	92.2	79.8	
1.159	34.77	1.110	2220	298	14.1	6.4	14	12.4	86.6	95.8	83.0	

^aNominal propellant utilization, 89%.

^bNominal losses, 273 eV/ion.

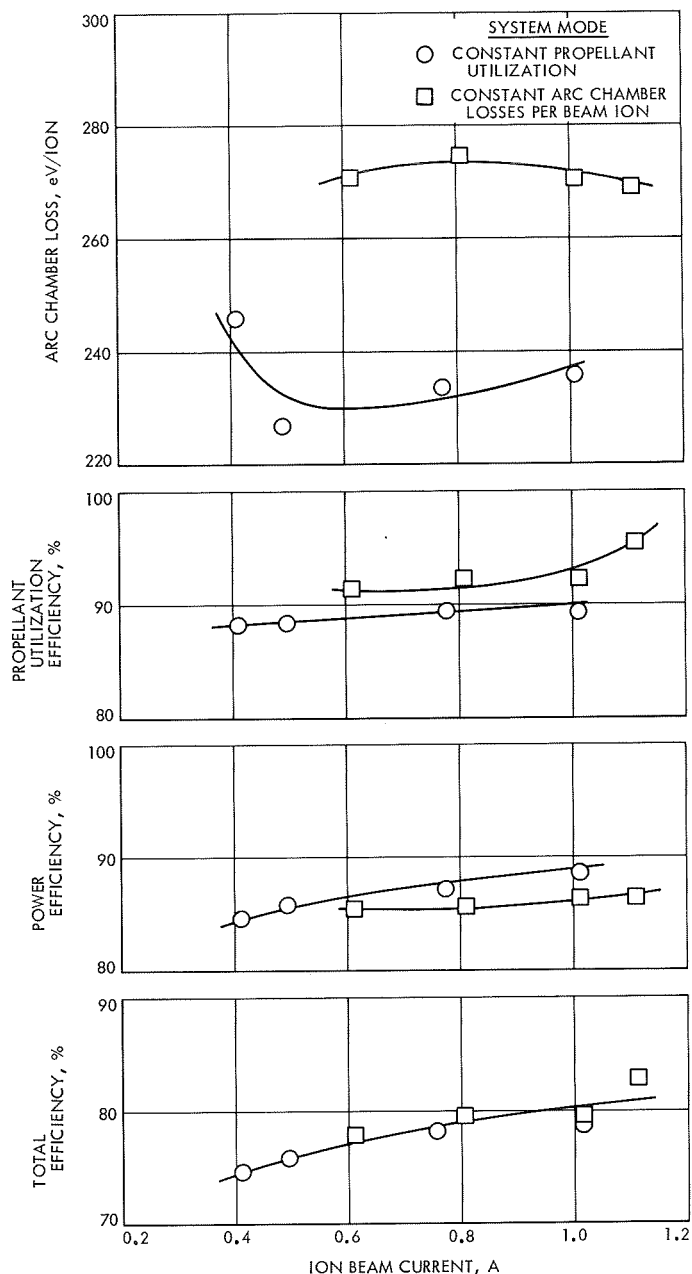


Fig. 2. Thruster efficiency and arc chamber losses for two throttling modes

corresponding to a propellant utilization of about 83%. This problem will continue to be investigated. Slight changes in the level of the magnetic field have proved to be a partial answer for the type of problem with the oxide-cathode thruster and therefore this approach will be pursued.

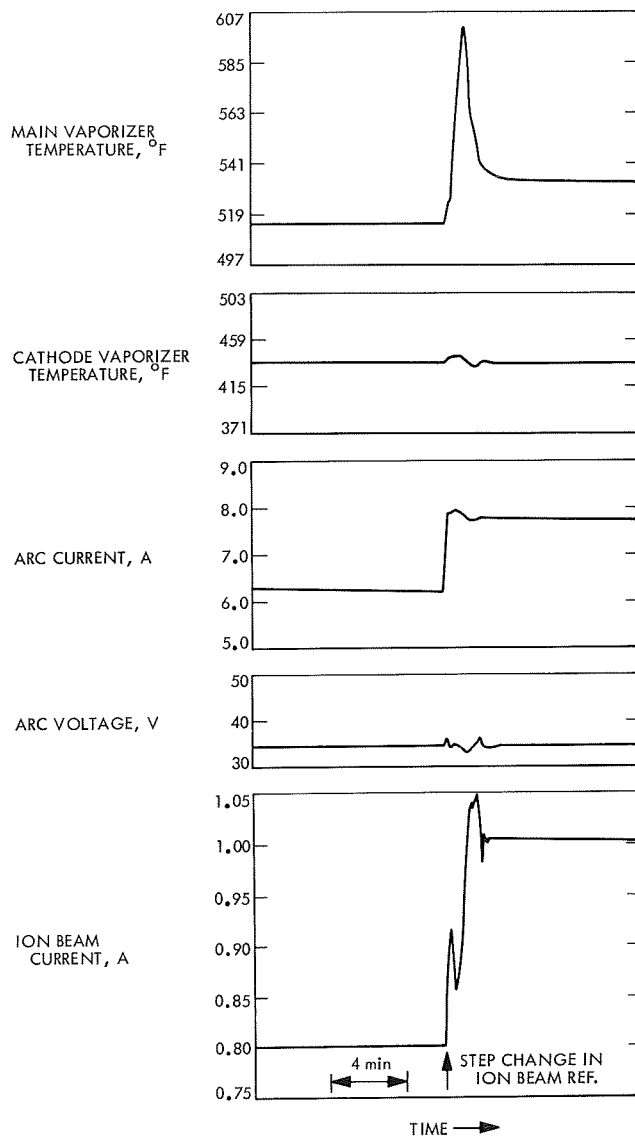
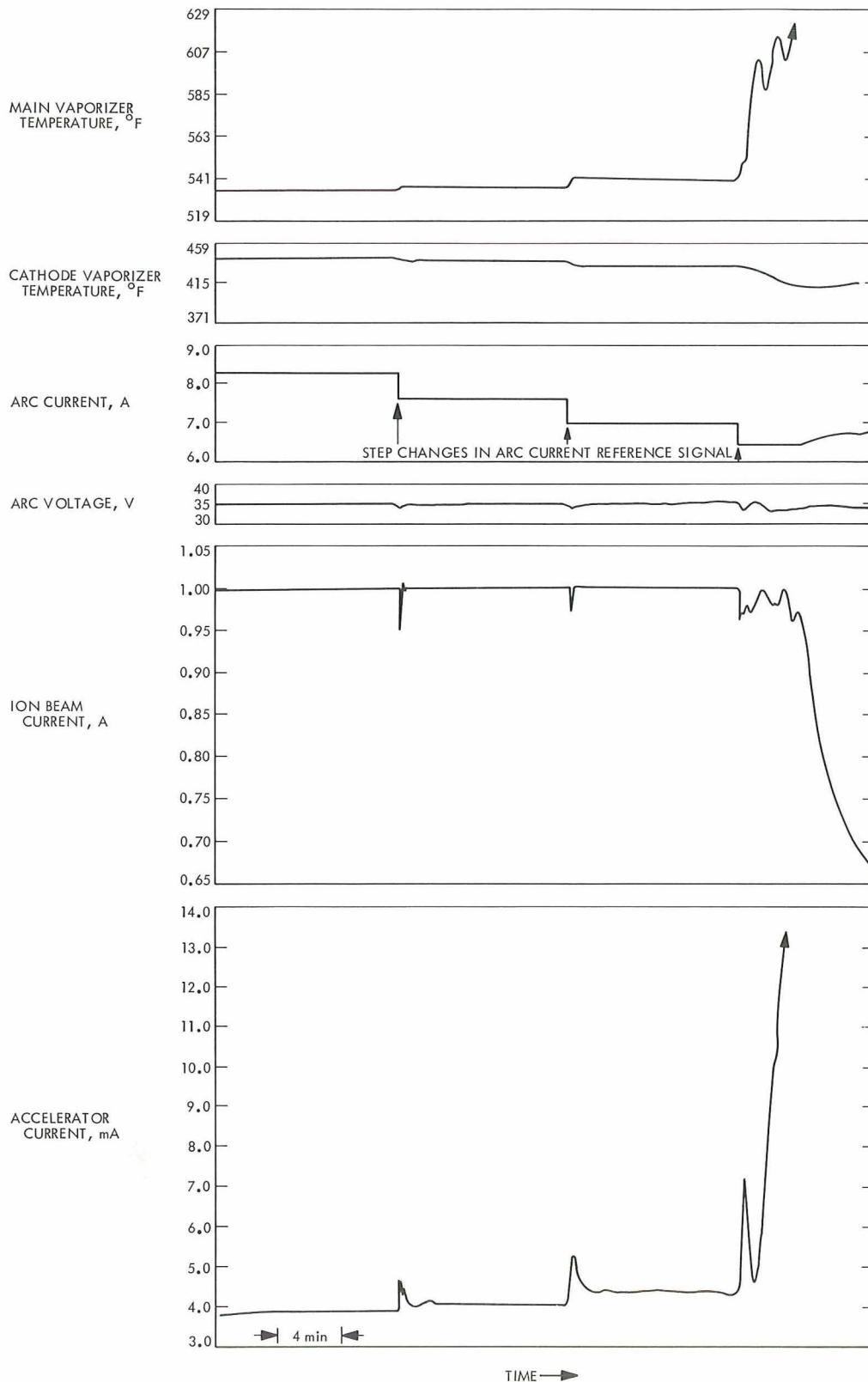


Fig. 3. System response to a step change in output power reference

References

1. Masek, T. D., "Plasma Properties and Performance of Mercury Ion Thrusters," Paper 69-256, presented at the Seventh AIAA Electric Propulsion Conference, Williamsburg, Va., Mar. 1969.
2. Mueller, P. A., and Pawlik, E. V., "Control Analysis of an Ion Thruster with Programmed Thrust," Paper 69-239, presented at the Seventh AIAA Electric Propulsion Conference, Williamsburg, Va., Mar. 1969.
3. Bechtel, R. T., "Performance and Control of a 30-Centimeter Diameter Low Impulse Kaufman Thruster," Paper 69-238, presented at the Seventh AIAA Electric Propulsion Conference, Williamsburg, Va., 1969.



**Fig. 4. System response to decreasing arc current set points
(function generator disconnected)**

XIX. Liquid Propulsion

PROPULSION DIVISION

A. Material Compatibility, *L. R. Toth and O. F. Keller*

The effects of long-term storage for the earth-storable liquid propellants hydrazine (N_2H_4) and nitrogen tetroxide (N_2O_4) were initially discussed in SPS 37-46, Vol. IV, pp. 162-172. Phase II incorporates some basic improvements, such as improved techniques for the measurement of test-capsule pressures. The stainless steel Bourdon-tube-type gage, utilized in the initial program, as reported in SPS 37-17, Vol. IV, p. 198, has been replaced by an externally mounted transducer to prevent an interaction between the Bourdon gage material and the propellant. In order to allow for a greater number of test-specimen/capsules to be tested simultaneously, automatic data acquisition and automatic data processing have been incorporated.

The specification-grade propellants selected for testing under the Phase II program have been divided into fuels and oxidizers as follows:

Fuels:

Hydrazine (N_2H_4)

Neat monomethylhydrazine (CH_3NHNH_2)

Hydrazine-hydrazine nitrate mix ($\text{N}_2\text{H}_4\text{-N}_2\text{H}_5\text{NO}_3$)

Oxidizers:

Inhibited nitrogen tetroxide (N_2O_4)

The propellants have been analyzed before storage testing for characterization and comparison with specification requirements. Wherever possible, all propellant samples of a given type have been taken from the same propellant drum in order to reduce effects from variations in propellant constituents. During testing, 20 cm³ of propellant will be tested in each specimen/capsule, as shown in Fig. 1.

The specimens selected for testing under the Phase II program were divided into various classes of materials, such as: stainless steels, aluminums, titaniums, plastics, and others.

These were further divided according to specimen configuration as follows: bimetal contact, bimetal separated, stressed, welded, brazed, coated, and plated.

One of the program goals is to maintain a high degree of control over all phases of specimen preparation and testing, including dimensional control and the use of certified materials. Test specimens are typically 3 in. long, 1/2 in. wide, and 1/32 in. thick. The specimens are partially immersed in propellant as shown in Fig. 1.

Post-test analyses of specimens planned at this time will include determination of changes in physical condition such as changes in weight, color, and surface condition. Subtle effects, such as intergranular corrosion, will also be investigated.

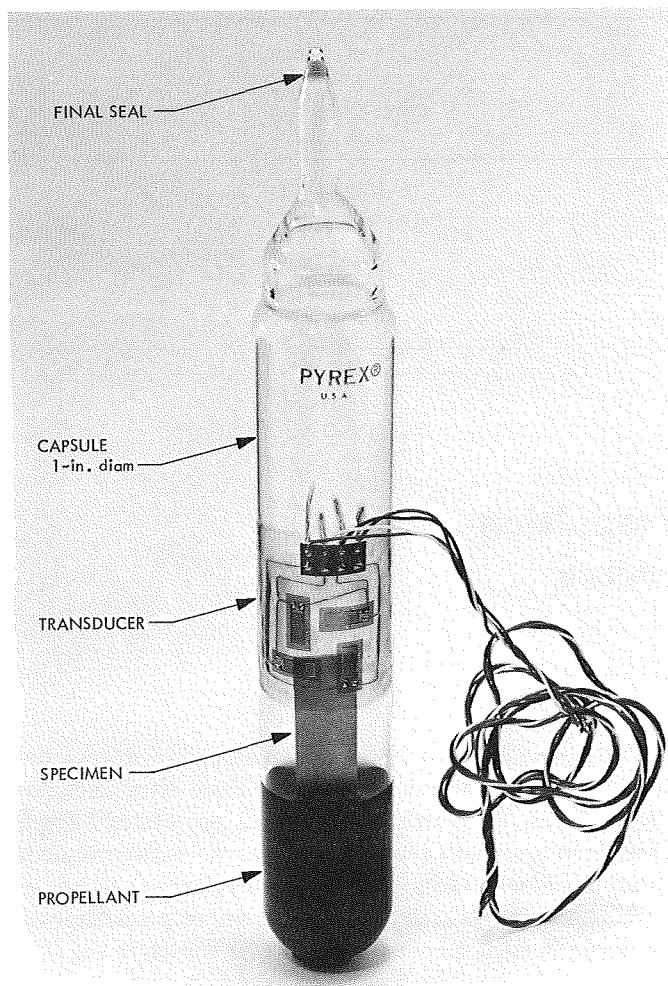


Fig. 1. Material compatibility specimen/capsule

The capsules are made of commercial grade pyrex, a borosilicate glass, which has been carefully annealed to avoid internal stresses that could affect transducer operation. A transition section, also of pyrex glass, has been fused to the capsule opening and finally, at the top, a hermetic seal has been made by glass fusion to contain any pressure generated within the capsule during the test.

Since glass is noncrystalline in nature, the working stress within the glass has been kept low, on the order of 875 psi, to avoid any tendency for glass creep to occur at the ambient test temperature of 110°F and the maximum test termination pressure of 50 psia that has been assumed for this program.

A transducer, consisting of four strain gages, connected as a Wheatstone bridge, is carefully bonded to each glass capsule for the purpose of both sensing capsule internal

pressures and converting these pressures to electrical signals which can be automatically recorded on data cards. The transducer generates an electrical signal of about $2 \mu\text{V}$ for each micro-inch expansion or contraction of the capsule. A transducer output of $4 \mu\text{V}$ is approximately equivalent to 1 psi differential pressure across the capsule wall. Because of the sensitive nature of the transducer, the proper processing and curing of the bonding material is important.

The automatic feature of the data acquisition system is also important because of the repetitive nature of the pressure readings, taken over periods of 2 yr or more, and because of the relatively large number of specimen/capsules that may be tested at any one time. Each test fixture has the capacity for testing approximately 900 test-specimen/capsules simultaneously. One fixture has been provided for fuel testing; the other for oxidizer testing.

Some environmental testing of hydrazine specimen/capsules at 110°F has been started at the material compatibility test building located at the Edwards Test Station; however, more work is required to make the automatic data acquisition system operational and to initiate automatic data processing.

B. Advanced Injector Fabrication Techniques, R. W. Riebling

1. Introduction

The sensitivity of the discharge coefficients of rocket motor injection orifices to the methods by which they are made has been experimentally demonstrated previously (Refs. 1 and 2) and is related to the effects of the fabrication process on the entrance region and interior surface of the orifice. The theoretical relations between wall roughness, entrance conditions, and discharge coefficient are summarized in Ref. 3.

One relatively new technique being applied to injector orifice fabrication is electroforming. A brief description of this process in general may be found in Ref. 4, while its recent application to rocket injectors is outlined in Ref. 5. The Liquid Propulsion Section is managing advanced technology efforts in the field of electroformed injectors for the National Aeronautics and Space Administration.¹ This article describes the results of a brief study of the visual characteristics of state-of-the-art electroformed injection orifices carried out in support of that management function.

¹Contract NAS 7-691, Camin Laboratories, Inc., Brooklyn, N. Y.

2. Effect of Fabrication Process on Appearance of Interior Surfaces

Orifices (0.029-in. nominal diameter) fabricated in Type A nickel by three different methods were longitudinally sectioned and their interior surfaces examined at approximately 500 \times magnification with a scanning electron microscope. Representative results are shown in Fig. 2. The twist-drilled orifice (Fig. 2a) exhibits helical grooves on its interior surface, while the electric-discharge-machined orifice (Fig. 2b) has a highly pitted interior surface, probably due to the local melting and removal of material by the electric spark. No attempt at quantitative surface roughness measurement was made, but it is readily apparent from the photographs that the surface finish, and therefore the friction factor of these orifices, is markedly influenced by the fabrication process. An orifice electroformed around a glass tube mandrel (subsequently removed by chemical dissolution) shows a much smoother interior surface, which most likely corresponds to a lower friction factor and higher discharge coefficient, all else being the same (Fig. 2c).

The electroforming process tends to faithfully reproduce the surface characteristics of the mandrel on the electroform itself, as shown in Fig. 3. In this case, the mandrel was a conventionally machined stainless steel rod, subsequently mechanically separated from the electroform by axial pulling. The tooling marks present on the mandrel have been reproduced on the inside of the orifice, and several large axially oriented grooves have resulted from mechanical removal of the mandrel. The appearance of the inside of this orifice is not strikingly different from that of the drilled hole (Fig. 2a). Comparison of Figs. 2c and 3 would tend to indicate that electroforming can produce interior surface finishes apparently superior to those attainable by other techniques when smooth-walled mandrels and chemical mandrel removal procedures are employed. Such potential is easily compromised when the standards of mandrel fabrication and removal are relaxed.

3. Electroformed Micro-Orifices

So-called "micro-orifice" injectors are described in Refs. 3 and 6. These make use of a great many orifices, each with a flow area on the order of 10^{-4} to 10^{-5} in.², arranged in close proximity across the face of the injector. Their advantages include high combustion efficiency as a result of improved atomization and more intimate propellant mixing, and enhanced injector face cooling because of a transpiration effect. However, micro-orifice injectors have been somewhat limited by current fabrication techniques that form them from brazed stacks of thin metal

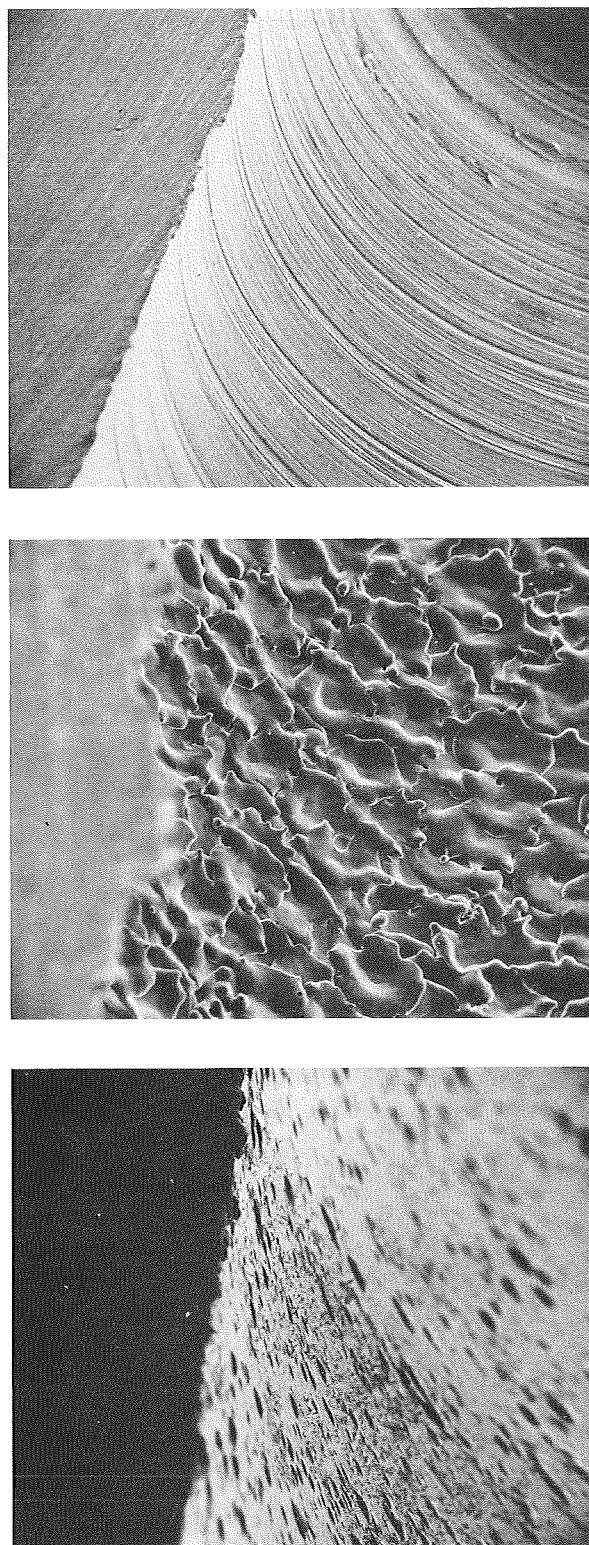


Fig. 2. Cross-sectional views of 0.029-in.-diam-injector orifices (Type A nickel) fabricated by various methods: (a) twist-drilled, (b) electric-discharge machined, (c) electroformed, using glass tube mandrel

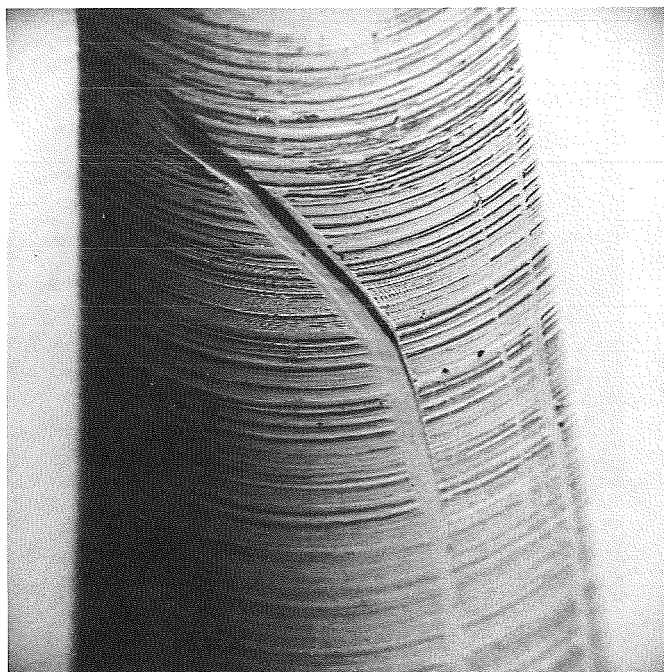


Fig. 3. Scanning electron microscope cross-sectional view of 0.029-in.-diam orifice electroformed from Type A nickel using machined stainless steel mandrel

platelets containing etched flow channels. Such platelet stacks are excellent heat exchangers, which often leads to an inability to exert full control over propellant injection temperatures. For these reasons, and also to more fully assess the capabilities of the electroforming process, the technology of electroformed micro-orifices is being developed as part of the previously cited JPL-managed NASA contract.

These orifices are formed by electroforming nickel from a sulfamate bath around precision-diameter hollow glass tubes, typically 0.003-in. in diameter. The exterior surface of the glass tubes, which are formed by a drawing process, is extremely smooth, as illustrated by the 8500 \times scanning electron microscope photograph of Fig. 4a. If previous electroforming technology is scalable downward, it would be expected that the interior surfaces of micro-orifices so formed would be transfers of the glass surfaces, and therefore equally smooth.

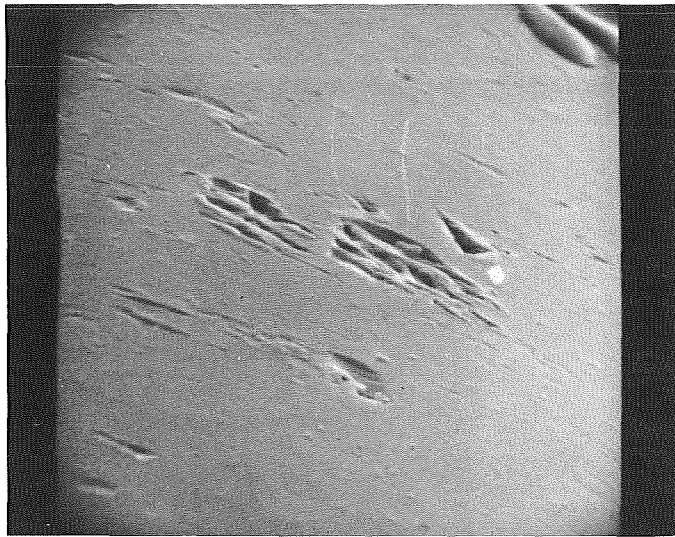
The end view of an 0.003-in.-diam electroformed micro-orifice is shown in Fig. 4b. These holes are typically well-formed and nearly circular, although at the present state-of-the-art some bell-mouthing of the exit frequently occurs. This may be minimized or eliminated by

electropolishing the ends of the orifices subsequent to electroforming.

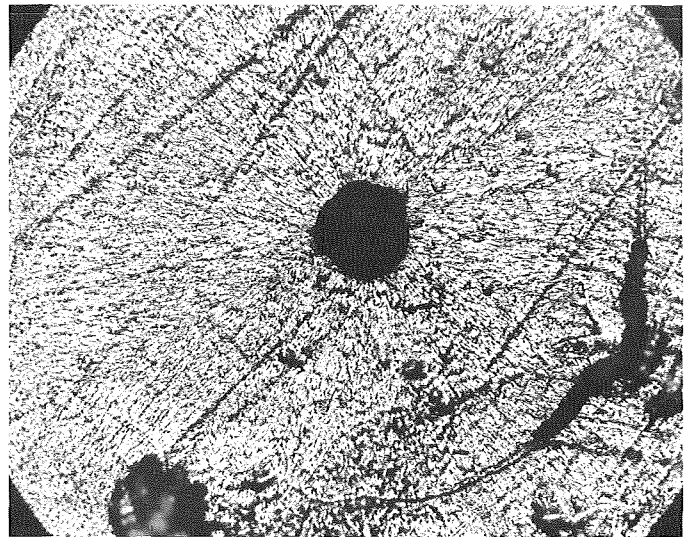
Longitudinal sections (at 500 \times magnification on scanning electron microscope) through 0.003-in.-diam orifices electroformed in nickel are shown in Figs. 4c and 4d. Both orifices were electroformed in an identical manner but cleaned differently. The glass capillary was removed from the orifice of Fig. 4c by simple immersion in hydrofluoric acid, while glass removal was effected by immersion with ultrasonic agitation, followed by flowthrough of the acid, in the case of the orifice of Fig. 4d. It may be seen that simple immersion is not effective for removing the glass pins, as evidenced by the rough glass residue adhering to the inside of the hole of Fig. 4c. Ultrasonic agitation and throughput of the chemical etchant, however, proved to be highly successful. No glass residue can be detected in photomicrographs such as Fig. 4d. Complete removal of the glass exposes a number of interfacial pits on the inside surface. As these pits are only about 0.0005 in. in length, they do not pose a serious threat in terms of providing lodging places for particulate contaminants or markedly affecting friction factors. Such state-of-the-art micro-orifices are therefore deemed acceptable for application to rocket engine injectors. Nevertheless, continuing research is expected to provide electroforming criteria which will eliminate such pits entirely.

References

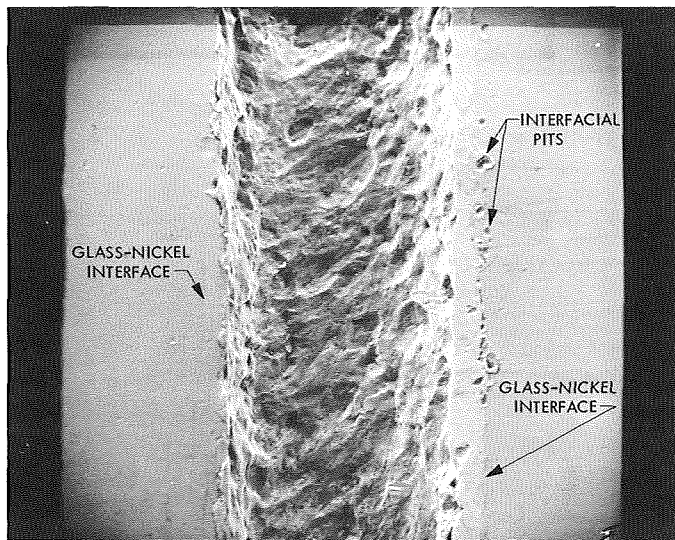
1. *Injector Orifice Study, Apollo Service Propulsion System*, Phase I Report, Contract NAS 9-6925, Aerojet-General Corp., Sacramento, Calif., Oct. 27, 1967.
2. *Chamber Technology for Space-Storable Propellants*, Report No. PES-68-2, Contract NAS 7-304, Rocketdyne, Canoga Park, Calif., Mar. 6, 1968.
3. Riebling, R. W., and Powell, W. B., *The Hydraulic Characteristics of Flow Through Miniature Slot Orifices*, Technical Report 32-1397, Jet Propulsion Laboratory, Pasadena, Calif., Sept. 15, 1969.
4. Kura, J. G., et al., *The Making of Nickel and Nickel-Alloy Shapes by Casting, Powder Metallurgy, Electroforming, Chemical Vapor Deposition, and Metal Spraying*, Technical Memorandum X-53430, George C. Marshall Space Flight Center, Huntsville, Ala., Oct. 1965.
5. Hammer, S. S., and Cacha, Z., "Development of Advanced Electroforming Techniques for Injector and Composite Thrust Chamber Fabrication," Paper 69-583, AIAA 5th Propulsion Joint Specialist Conference, U.S. Air Force Academy, Colorado, June 9-13, 1969.
6. Duckering, R. E., *Advanced Injectors for Space Storable Propellants*, Report 713-F, Aerojet-General Corp., Sacramento, Calif., Oct. 1969.



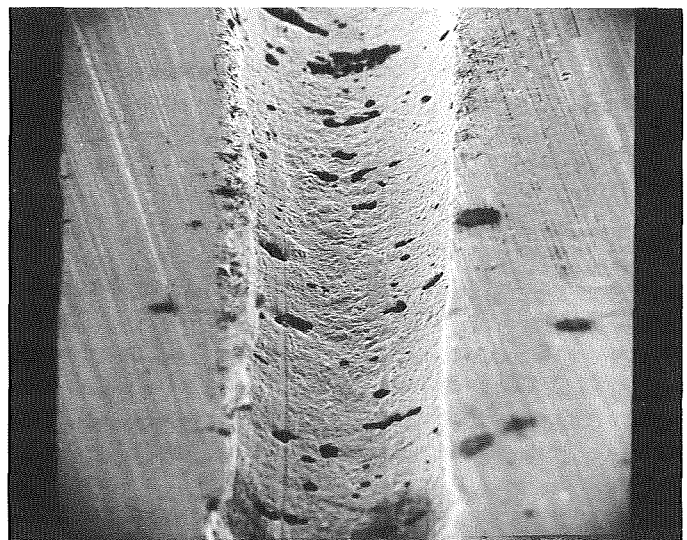
(a) EXTERIOR SURFACE OF 0.003-in. O. D. GLASS CAPILLARY TUBE MANDREL



(b) END VIEW OF 0.003-in. diam ORIFICE ELECTROFORMED IN TYPE A NICKEL



(c) LONGITUDINAL SECTION OF 0.003-in. diam ORIFICE ELECTROFORMED IN TYPE A NICKEL, IMMERSION CLEANING



(d) LONGITUDINAL SECTION OF 0.003-in. diam ORIFICE ELECTROFORMED IN TYPE A NICKEL, FLOW-THROUGH CLEANING

Fig. 4. Visual characteristics of electroformed micro-orifices and mandrels

XX. Systems Analysis Research

MISSION ANALYSIS DIVISION

A. Relativistic Solar Perturbations on the Moon in the Brans-Dicke Theory, J. D. Anderson

With the placement of corner reflectors on the moon by *Apollo* astronauts, thereby making possible laser ranging to the moon to an accuracy of a few centimeters, and with the possibility of the placement of active radio transponders during future lunar landings, it is important to consider post-Newtonian terms in the motion of the moon in considerable detail. Laser and transponding ranging make possible the measurement of both secular relativistic terms, for example, the familiar DeSitter-Schouten effect (Ref. 1), and also periodic terms in the moon's motion. In this article the relativistic perturbative acceleration on the moon by the sun is derived in an approximate form which will allow an exploration of the magnitude of the secular and periodic terms. Then it should be possible to determine to what degree lunar ranging can separate the post-Newtonian solar terms from other lunar motions and to determine the expected accuracy in measurements of the magnitude of these relativistic terms.

In a recent paper by Estabrook (Ref. 2), the post-Newtonian n -body equations of motion are derived for the Brans-Dicke theory. The acceleration of the i th par-

ticle in a system of n particles is given by the following expression:

$$\begin{aligned} \ddot{\mathbf{r}}_i = & \sum_{j \neq i} \frac{\mu_j (\mathbf{r}_j - \mathbf{r}_i)}{r_{ij}^3} \left\{ 1 - \frac{1}{c^2} \frac{6 + 4\omega}{2 + \omega} \sum_{l \neq i} \frac{\mu_l}{r_{il}} \right. \\ & - \frac{1}{c^2} \sum_{k \neq j} \frac{\mu_k}{r_{jk}} + \frac{1}{c^2} \frac{1 + \omega}{2 + \omega} v_i^2 + \frac{1}{c^2} \frac{3 + 2\omega}{2 + \omega} v_j^2 \\ & - \frac{2}{c^2} \frac{3 + 2\omega}{2 + \omega} \dot{\mathbf{r}}_i \cdot \dot{\mathbf{r}}_j - \frac{3}{2c^2} \left[\frac{(\mathbf{r}_j - \mathbf{r}_i) \cdot \dot{\mathbf{r}}_j}{r_{ij}} \right]^2 \\ & + \frac{1}{2c^2} (\mathbf{r}_j - \mathbf{r}_i) \cdot \ddot{\mathbf{r}}_j \left. \right\} - \frac{1}{c^2} \sum_{j \neq i} \frac{\mu_j}{r_{ij}^3} (\mathbf{r}_j - \mathbf{r}_i) \\ & \cdot \left(\frac{4 + 3\omega}{2 + \omega} \dot{\mathbf{r}}_j - \frac{6 + 4\omega}{2 + \omega} \dot{\mathbf{r}}_i \right) (\dot{\mathbf{r}}_j - \dot{\mathbf{r}}_i) \\ & + \frac{1}{2c^2} \frac{10 + 7\omega}{2 + \omega} \sum_{j \neq i} \frac{\mu_j \ddot{\mathbf{r}}_j}{r_{ij}} \end{aligned} \quad (1)$$

where ω is the coupling constant of the scalar field.

In the three-body problem for the sun, earth, and moon, the origin of coordinates can be taken at the sun, and the relativistic solar perturbations on the motion of the moon can be approximated by setting the masses of the earth and moon equal to zero. Then the heliocentric equations of motion for both the earth and moon will contain only the solar terms in the summation on j in Eq. (1). The gravitational constant μ_j can be replaced by

$$\mu_j = mc^2 \quad (2)$$

where m , one-half the gravitational radius of the sun, is equal to 1.4766 km. Further, because the origin of coordinates is at the sun, both \mathbf{r}_j and $\dot{\mathbf{r}}_j$ are zero. As a result, the expressions for the heliocentric accelerations of the earth and moon, $\ddot{\mathbf{r}}_\oplus$ and $\ddot{\mathbf{r}}_\epsilon$, respectively, are greatly simplified and are perfectly symmetrical. For the earth the solar perturbation is

$$\begin{aligned} \ddot{\mathbf{r}}_\oplus = & \frac{m\mathbf{r}_\oplus}{r_\oplus^3} \left[\frac{6 + 4\omega}{2 + \omega} \frac{mc^2}{r_\oplus} - \frac{1 + \omega}{2 + \omega} v_\oplus^2 \right] \\ & + \frac{m\dot{\mathbf{r}}_\oplus}{r_\oplus^3} \left(\frac{6 + 4\omega}{2 + \omega} \mathbf{r}_\oplus \cdot \dot{\mathbf{r}}_\oplus \right) \end{aligned} \quad (3)$$

and the perturbation on the moon is exactly the same with the subscript \oplus replaced by ϵ .

The solar perturbation on the geocentric motion of the moon can be found by differencing the heliocentric accelerations

$$\ddot{\mathbf{r}}_{\oplus\epsilon} = \ddot{\mathbf{r}}_\epsilon - \ddot{\mathbf{r}}_\oplus \quad (4)$$

First, however, it is advantageous to assume that $\mathbf{r}_\oplus = \mathbf{r}_\epsilon = \mathbf{R}$ and to neglect the radial velocity $\dot{\mathbf{r}}_\oplus$ of the earth. Then the difference of accelerations is

$$\ddot{\mathbf{r}}_{\oplus\epsilon} = \frac{m\dot{\mathbf{r}}_\epsilon}{R^3} \left(\frac{6 + 4\omega}{2 + \omega} \mathbf{R} \cdot \dot{\mathbf{r}}_\epsilon \right) + \frac{m\mathbf{R}}{R^3} \frac{1 + \omega}{2 + \omega} (v_\oplus^2 - v_\epsilon^2) \quad (5)$$

However,

$$\dot{\mathbf{r}}_\epsilon = \dot{\mathbf{r}}_\oplus + \dot{\mathbf{r}}_{\oplus\epsilon} \quad (6)$$

and

$$v_\epsilon^2 = \dot{\mathbf{r}}_\epsilon \cdot \dot{\mathbf{r}}_\epsilon \approx v_\oplus^2 + 2\dot{\mathbf{R}} \cdot \dot{\mathbf{r}}_{\oplus\epsilon} \quad (7)$$

Also,

$$\mathbf{R} \cdot \dot{\mathbf{r}}_\epsilon = \mathbf{R} \cdot \dot{\mathbf{r}}_{\oplus\epsilon} \quad (8)$$

Therefore, to the first order in the geocentric velocity of the moon, the relativistic solar perturbation is

$$\ddot{\mathbf{r}}_{\oplus\epsilon} = \frac{2m}{R^3} \left\{ \frac{3 + 2\omega}{2 + \omega} (\mathbf{R} \cdot \dot{\mathbf{r}}_{\oplus\epsilon}) \dot{\mathbf{R}} - \frac{1 + \omega}{2 + \omega} (\dot{\mathbf{R}} \cdot \dot{\mathbf{r}}_{\oplus\epsilon}) \mathbf{R} \right\} \quad (9)$$

This equation has been derived for the Einstein case ($\omega \rightarrow \infty$) in a different form by Tausner (Ref. 3). In terms of the Robertson parameter γ used in an earlier article (SPS 37-50, Vol. III, pp. 39-47), the acceleration is given by

$$\ddot{\mathbf{r}}_{\oplus\epsilon} = \frac{2m}{R^3} \{ (1 + \gamma) (\mathbf{R} \cdot \dot{\mathbf{r}}_{\oplus\epsilon}) \dot{\mathbf{R}} - \gamma (\dot{\mathbf{R}} \cdot \dot{\mathbf{r}}_{\oplus\epsilon}) \mathbf{R} \} \quad (10)$$

References

1. Eddington, A. S., *The Mathematical Theory of Relativity*, Section 44. Cambridge University Press, London, 1924.
2. Estabrook, F. B., *Astrophys. J.*, Vol. 158, p. 81, 1969.
3. Tausner, M. J., *General Relativity and Its Effects on Planetary Orbits and Interplanetary Observations*, MIT Technical Report 425, Massachusetts Institute of Technology, Cambridge, Mass., Oct. 7, 1966.

B. Propagation of Statistical Moments in Nonlinear Systems and the Application to a Planetary Orbiter, J. F. Jordan

1. Introduction

Often it is desired to know the solution in the statistical sense of a nonlinear differential equation governing the behavior of an N dimensional state $x_i, i = 1, 2, \dots, N$, of the form

$$\dot{x}_i = f_i(x, t) \quad (1)$$

where the initial state at $t = 0$, $x_i(0)$, is a random variable vector with known joint probability density function $g(x)$ or, equivalently, known statistical moments

$$\left. \begin{aligned} E[x_{i_0}] &= \mu_i \\ E[(x_{i_0} - \mu_i)(x_{j_0} - \mu_j)] &= P_{ij_0} \\ E[(x_{i_0} - \mu_i)(x_{j_0} - \mu_j)(x_{k_0} - \mu_k)] &= P_{ijk_0} \\ \text{etc.} \end{aligned} \right\} \quad (2)$$

The solution of differential equation (1) takes the form of the joint density function of $x_i(t)$ or the time varying moments

$$\left. \begin{aligned} \bar{x}_i(t) &= E(x_i(t)) \\ P_{ij}(t) &= E[(x_i(t) - \bar{x}_i(t))(x_j(t) - \bar{x}_j(t))] \\ \text{etc.} \end{aligned} \right\} \quad (3)$$

Since, when f_i is a nonlinear function of x , the time varying density is often difficult to solve for, it is desired to find equations whose solutions yield the statistical moments defined in Eq. (3). This procedure has practical advantages since often the mean \bar{x}_i and the variance P_{ij} are of more interest to the analyst than the density function itself. The problem of propagating statistical moments has been the subject of several papers, including Ref. 1.

2. The Differential Equation for the Statistical Moments

Taking the expected value of Eq. (1) and introducing the commutative property of the expected value and the differentiation operations leads to

$$\frac{d}{dt} \bar{x}_i = \bar{f}_i(x, t) \quad (4)$$

Assuming the existence of the time varying mean \bar{x}_i , expand the function f_i in a Taylor series about \bar{x}_i to yield

$$f_i(x, t) = f_i(\bar{x}, t) + f_{ix_j} \delta x_j + \frac{1}{2} f_{ix_j x_k} \delta x_j \delta x_k + \dots \quad (5)$$

where

$$\begin{aligned} \delta x_i &= x_i - \bar{x}_i \\ f_{ix_j} &= \frac{\partial f_i}{\partial x_j}(\bar{x}, t), \text{ etc., } \dots \end{aligned}$$

and the repeated subscript denotes summation.

Substitution of Eq. (5) into Eq. (4) and carrying out the expected value operation yields the following differential equation:

$$\dot{\bar{x}}_i = f_i(\bar{x}, t) + \frac{1}{2} f_{ix_j x_k} P_{jk} + \frac{1}{3!} f_{ix_j x_k x_l} P_{jkl} + \dots \quad (6)$$

Differentiation of P_{ij} yields, from the definition,

$$\dot{P}_{ij} = E[\dot{\delta x}_i \delta x_j] + E[\delta x_i \dot{\delta x}_j] \quad (7)$$

But from the subtraction of Eq. (6) from Eq. (1), the following equation is formed:

$$\delta \dot{x}_i = f_{ix_j} \delta x_j + \frac{1}{2} f_{ix_j x_k} (\delta x_j \delta x_k - P_{jk}) + \dots \quad (8)$$

Substitution of Eq. (8) into Eq. (7) and carrying out the expected value operations then yields

$$\begin{aligned} \dot{P}_{ij} &= f_{ix_k} P_{kj} + P_{ik} f_{jx_k} + \frac{1}{2} f_{ix_k x_l} P_{klj} \\ &+ \frac{1}{2} P_{ikl} f_{jx_k x_l} + \dots \end{aligned} \quad (9)$$

Using the definition of P_{ijk} and the same reasoning, the following equation can be derived:

$$\dot{P}_{ijk} = f_{ix_l} P_{ljk} + P_{ikl} f_{jx_l} + P_{ijl} f_{kx_l} + \dots \quad (10)$$

and so on for the higher moments P_{ijkl} , etc., \dots

In many problems the density function for the initial state is gaussian or near gaussian, such that the higher moments are small compared with \bar{x}_i and P_{ij} . Also, in cases of slight nonlinearities in the equation of motion, the higher derivatives of f_i with respect to x are small. When either of these cases occur, the following approximate equations for the moment differential equations are suggested:

$$\left. \begin{aligned} \dot{\bar{x}}_i &= f_i(\bar{x}, t) + \frac{1}{2} f_{ix_j x_k} P_{jk} \\ \dot{P}_{ij} &= f_{ix_k} P_{kj} + P_{ik} f_{jx_k} \end{aligned} \right\} \quad (11)$$

with higher moments set to zero.

3. An Analytical Example

Consider the one-dimensional nonlinear differential equation

$$\dot{x} = x^2 \quad (12)$$

where the density function of the initial state x_0 is given by

$$g(x_0) = \frac{1}{2a}, \quad \mu - a < x_0 < \mu + a$$

The mean and variance of x_0 are easily computed by the following integrals:

$$\left. \begin{aligned} \bar{x}_0 &= E(x_0) = \int_{\mu-a}^{\mu+a} \frac{x_0}{2a} dx_0 = \mu \\ P_0 &= E[(x_0 - \bar{x}_0)^2] = \int_{\mu-a}^{\mu+a} \frac{(x_0 - \mu)^2}{2a} dx_0 = \frac{a^2}{3} \end{aligned} \right\} \quad (13)$$

From the analytical solution of Eq. (12), i.e.,

$$x(t) = \frac{x_0}{1 - x_0 t} \quad (14)$$

the density function of $x(t)$ can be found from the use of the standard change of variable technique:

$$g(x(t)) = \frac{1}{2a(1 + xt)^2} \quad (15)$$

The mean $\bar{x}(t)$ can then be computed by integration and written in the following series form:

$$\bar{x}(t) = \frac{\mu}{1 - \mu t} + \frac{1}{3} \frac{a^2 t}{(1 - \mu t)^3} + \frac{1}{5} \frac{a^4 t^3}{(1 - \mu t)^5} + \dots \quad (16)$$

The differential equations for $\bar{x}(t)$ and $P(t)$ are computed from Eq. (11), where $f(x, t) = x^2$. They are as follows:

$$\left. \begin{aligned} \dot{\bar{x}} &= \bar{x}^2 + P \\ \dot{P} &= 4\bar{x}P \end{aligned} \right\} \quad (17)$$

The numerical solution of Eq. (17) is compared with the analytical solution, Eq. (16) and the first-order solution, $\bar{x}(t) = \mu/(1 - \mu t)$, in Fig. 1. Note that the solution to Eq. (17) offers a much better approximation to the true value of $\bar{x}(t)$ than does the first-order solution.

4. The Motion of an Orbiting Spacecraft

Often it is desired to know the time history of the mean state and the variance of the errors in the state of an orbiting spacecraft if these moments are known at some epoch from either *a priori* or observational information. Mapping of these quantities is useful if updated best estimates (conditional mean values) are desired or if long range prediction studies are needed. Consider the motion of a probe in elliptical motion around a planetary body.

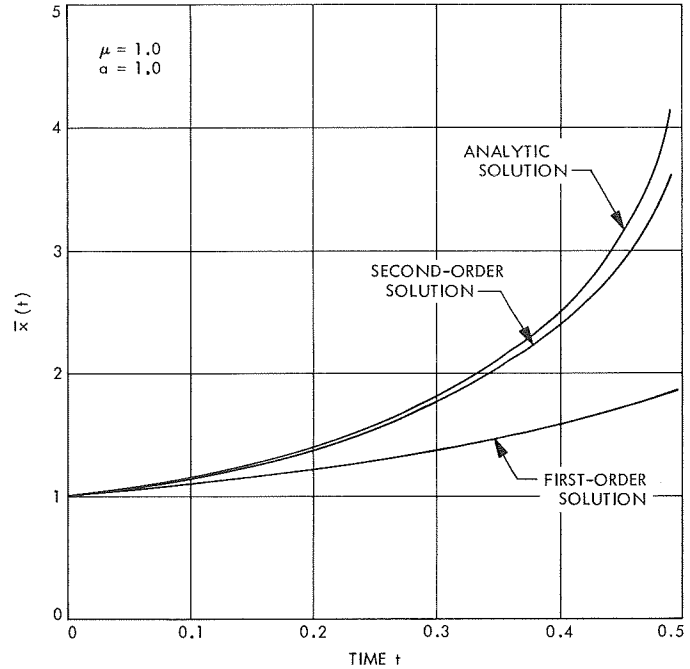


Fig. 1. One-dimensional mean state vs time (analytic example)

The equations of motion for the cartesian state (\mathbf{v}, \mathbf{x}) are given by

$$\left. \begin{aligned} \dot{\mathbf{v}} &= \frac{\mu}{r^3} \mathbf{x} \\ \dot{\mathbf{x}} &= \mathbf{v} \end{aligned} \right\} \quad (18)$$

The behavior of the mean velocity and position is studied for a 12-h Martian orbiter of the type being considered for the *Mariner Mars 1971* mission. The initial orbital elements relative to the cartesian system in Eq. (18) are given by

$$\begin{aligned} a &= 12651 \text{ km} \\ e &= 0.58 \\ T_p &= 0, 12 \text{ h}, 24 \text{ h, etc.} \\ \Omega &= 1.93 \text{ rad} \\ \omega &= 3.49 \text{ rad} \\ i &= 1.20 \text{ rad} \end{aligned}$$

and the *a priori* covariance matrix of the state uncertainties is diagonal with the following sigmas:

$$\sigma_{u_0} = \sigma_{v_0} = \sigma_{w_0} = 10 \text{ cm/s}$$

and

$$\sigma_{x_0} = \sigma_{y_0} = \sigma_{z_0} = 10 \text{ km}$$

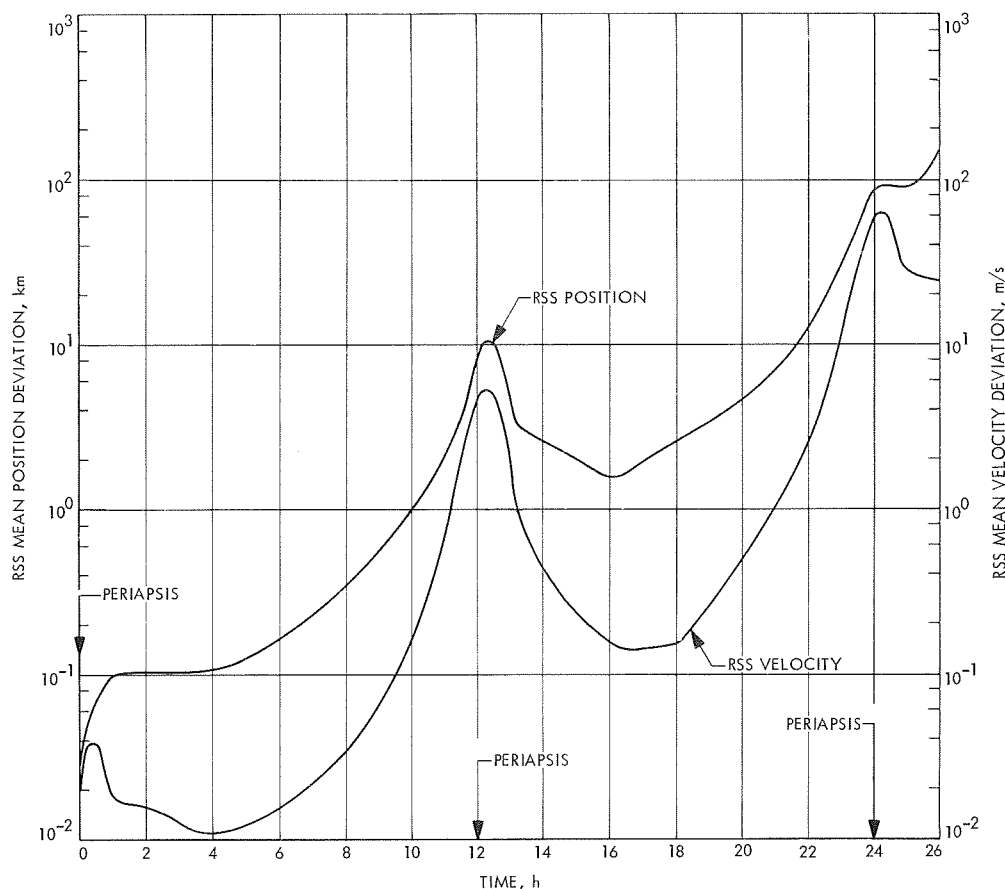


Fig. 2. Rss mean position and velocity deviation time histories (Mars orbiter)

The size of these sigmas is representative of the accuracy to which the state of a Martian orbiter can be estimated from one pass of earth-based doppler data.

The time histories of the root-sum-square (rss) mean deviations (difference between second-order and first-order solutions) of the orbiter velocity and position components are shown in Fig. 2 for a time span which includes two successive periapsis passages. The second-order solution for the rss mean position component differs from the deterministic or first-order solution by as much as 150 km at the second periapsis passage, while the velocity differs by 70 m/s.

It is evident from Fig. 2 that great care must be taken in mapping mean state values for an orbit or more, particularly if the mapped mean is to be combined with additional information from new measurements to form new estimates (conditional mean values) of the state vector. In many cases, the standard linear (first-order) mapping procedure may not be adequate and the second-order terms must be included. The application of the technique

presented here to the sequential filtering problem for a planetary orbiter will be the subject of a future article.

Reference

1. Cumming, I. G., "Derivation of the Moments of a Continuous Stochastic System," *Int. J. Control*, Vol. 5, No. 1, pp. 85-90, 1967.

C. The Statistics of the Solution of a Nonlinear Differential Equation With Random Initial Conditions, H. Lass

We consider the nonlinear differential equation

$$\frac{dx}{dt} = f(x) \quad (1)$$

with $x(0) = \alpha$ a random variable having a specified probability density function (PDF), say $p(\alpha)$. Under suitable conditions on $f(x)$, a unique solution of Eq. (1) exists, say

$x = \phi(t, \alpha)$. An important, but difficult, problem is to determine the various moments of the random variable x at any specified time t .

If x is known explicitly as a function of t and α , the PDF of x at time t is given by

$$\rho(x, t) = \int_{-\infty}^{\infty} \delta[x - \phi(t, \alpha)] p(\alpha) d\alpha \quad (2)$$

such that $\rho(x, t) dx$ is the probability at time t that the level of a solution of Eq. (1) lies in the range $(x, x + dx)$. The δ of Eq. (2) is the Dirac impulse function.

Equation (2) can be obtained as follows. Consider the following integral

$$\begin{aligned} F(x, t) &= \int_{-\infty}^x \int_{-\infty}^{\infty} \delta[\xi - \phi(t, \alpha)] p(\alpha) d\alpha d\xi \\ &= \int_{-\infty}^{\infty} \int_{-\infty}^x \delta[\xi - \phi(t, \alpha)] p(\alpha) d\xi d\alpha \end{aligned} \quad (3)$$

Now whenever $\phi(t, \alpha) < x$, the argument of the impulse function will vanish as we integrate on ξ over the range $(-\infty, x)$, thus yielding one unit to the value of the integral. The set of α for which $\phi(t, \alpha) < x$ is denoted by S , so that

$$F(x, t) = \int_S p(\alpha) d\alpha \quad (4)$$

We note that $F(x, t)$ is simply the distribution function associated with the random variable x , so that

$$\rho(x, t) = \frac{\partial F}{\partial x} = \int_{-\infty}^{\infty} \delta[x - \phi(t, \alpha)] p(\alpha) d\alpha \quad (5)$$

by virtue of Eq. (3).

By a similar reasoning, the joint PDF for $x = \phi(t, \alpha)$ and $v = \dot{x} = \dot{\phi}(t, \alpha) = \partial\phi/\partial t$ can be shown to be

$$\rho(x, v, t) = \int_{-\infty}^{\infty} \delta[x - \phi(t, \alpha)] \delta[v - \dot{\phi}(t, \alpha)] p(\alpha) d\alpha \quad (6)$$

The conditional PDF, $\rho(v, t|x)$ such that $\rho(v, t|x) dv$ is the probability that at time t the slope lies in the range

$(v, v + dv)$ given that $\phi(t, \alpha)$ has reached the level x , is from Bayes' formula

$$\rho(v, t|x) = \frac{\rho(x, v, t)}{\rho(x, t)} \quad (7)$$

Thus, the mean speed of all curves reaching the level x at time t is

$$\overline{v(x, t)} = \int_{-\infty}^{\infty} v \rho(v, t|x) dx = \frac{1}{\rho(x, t)} \int_{-\infty}^{\infty} v \rho(x, v, t) dv \quad (8)$$

From Eq. (6), we obtain

$$\begin{aligned} \rho(x, t) \overline{v(x, t)} &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} v \delta[x - \phi(t, \alpha)] \\ &\quad \times \delta[v - \dot{\phi}(t, \alpha)] p(\alpha) dv d\alpha \\ &= \int_{-\infty}^{\infty} \dot{\phi}(t, \alpha) \delta[x - \phi(t, \alpha)] p(\alpha) d\alpha \end{aligned} \quad (9)$$

A partial differentiation with respect to x yields

$$\frac{\partial}{\partial x} [\rho \overline{v}] = \int_{-\infty}^{\infty} \dot{\phi}(t, \alpha) \delta'[x - \phi(t, \alpha)] p(\alpha) d\alpha \quad (10)$$

From Eq. (5), we note that

$$\frac{\partial \rho}{\partial t} = - \int_{-\infty}^{\infty} \delta'[x - \phi(t, \alpha)] \dot{\phi}(t, \alpha) p(\alpha) d\alpha \quad (11)$$

Thus, we obtain the equation of continuity

$$\frac{\partial \rho}{\partial t} + \frac{\partial}{\partial x} (\rho \overline{v}) = 0 \quad (12)$$

Since $v = dx/dt = f(x)$ is the speed of any curve reaching the level at time t , of necessity $\overline{v} = f(x)$, so that

$$\frac{\partial \rho}{\partial t} = - \frac{\partial}{\partial x} [\rho f(x)] \quad (13)$$

Next we consider the k th central moment defined by

$$\mu_k(t) = \int_{-\infty}^{\infty} [x - \mu(t)]^k \rho(x, t) dx \quad (14)$$

with $\mu_0(t) \equiv 1$, $\mu_1(t) \equiv 0$, $\mu_2(t) = \sigma^2(t)$, the variance of x at any time t . The mean of x at any time t is

$$\mu(t) = \int_{-\infty}^{\infty} x \rho(x, t) dx \quad (15)$$

Now

$$\begin{aligned} \frac{d}{dt} \int_{-\infty}^{\infty} [x - \mu(t)]^k \rho(x, t) dx &= \int_{-\infty}^{\infty} [x - \mu(t)]^k \frac{\partial \rho}{\partial t} dx \\ &- k\mu'(t) \int_{-\infty}^{\infty} [x - \mu(t)]^{k-1} \rho(x, t) dx \end{aligned} \quad (16)$$

yielding

$$\frac{d\mu_k(t)}{dt} = -k\mu'(t) \mu_{k-1}(t) + \int_{-\infty}^{\infty} [x - \mu(t)]^k \frac{\partial \rho}{\partial t} dx \quad (17)$$

Next we multiply Eq. (13) by $[x - \mu(t)]^k$ and integrate on x over the range $(-\infty, \infty)$ to obtain

$$\int_{-\infty}^{\infty} [x - \mu(t)]^k \frac{\partial \rho}{\partial t} dx = - \int_{-\infty}^{\infty} [x - \mu(t)]^k \frac{\partial}{\partial x} [\rho f(x)] dx \quad (18)$$

Integrating by parts and using the fact that x is bounded in the time interval $(0, t)$ so that $\rho \equiv 0$ for $|x| > X$ yields

$$\int_{-\infty}^{\infty} [x - \mu(t)]^k \frac{\partial \rho}{\partial t} dx = k \int_{-\infty}^{\infty} [x - \mu(t)]^{k-1} \rho f(x) dx \quad (19)$$

Thus, Eq. (17) becomes

$$\frac{d\mu_k}{dt} = -k\mu'(t) \mu_{k-1}(t) + k \int_{-\infty}^{\infty} [x - \mu(t)]^{k-1} \rho f(x) dx \quad (20)$$

To express the integral of Eq. (20) in terms of the central moments, we expand $f(x)$ in a Taylor series about the mean $\mu(t)$. Thus,

$$f(x) = \sum_{r=0}^{\infty} \frac{f^{(r)}(\mu)(x - \mu)^r}{r!} \quad (21)$$

and Eq. (20) becomes

$$\frac{d\mu_k}{dt} = -k\mu'(t) \mu_{k-1} + k \sum_{r=0}^{\infty} \frac{1}{r!} f^{(r)}(\mu) \mu_{r+k-1} \quad (22)$$

for $k = 1, 2, 3, \dots$

Equation (20) points out immediately the difficulties involved in finding the central moments since $\dot{\mu}_k(t)$ depends linearly on the infinite set $\mu_{k-1}, \mu_k, \mu_{k+1}, \mu_{k+2}, \dots$, and depends nonlinearly on the mean $\mu(t)$.

The differential equation for $\mu(t)$ is obtained by setting $k = 1$ with $\mu_1(t) \equiv 0$, $\mu_0(t) \equiv 1$, so that

$$\frac{d\mu}{dt} = \sum_{r=0}^{\infty} \frac{1}{r!} f^{(r)}(\mu) \mu_r(t) \quad (23)$$

This equation can be obtained from $dx/dt = f(x)$ from the fact that the mean of dx/dt is $d\mu/dt$ with $\mu(t)$ the mean of x at time t . Thus, $d\mu/dt$ depends linearly on all the central moments but, in general, depends nonlinearly on $\mu(t)$. The special case $f(x) = ax$ yields $d\mu/dt = a\mu$, and $\mu(t) = \mu(0) e^{at}$.

Let us now apply Eq. (22) for $k = 1, 2, 3$ to the case $\dot{x} = x^2$, so that $f(x) = x^2$, $f'(\mu) = 2\mu$, $f''(\mu) = 2$, $f^{(n)}(\mu) = 0$ for $n > 2$. We obtain

$$\left. \begin{aligned} \frac{d\mu}{dt} &= \mu^2 + \sigma^2 \\ \frac{d\sigma^2}{dt} &= 4\mu\sigma^2 + 2\mu_3 \\ \frac{d\mu_3}{dt} &= 6\mu\mu_3 + 3(\mu_4 - \sigma^4) \end{aligned} \right\} \quad (24)$$

by virtue of $\dot{\mu} = \mu^2 + \sigma^2$.

Now suppose the PDF of α (initial condition) is symmetric about the mean so that $\mu_3(0) = 0$. Now, in general, $\mu_4 \geq \sigma^4$. If we assume that $\mu_4 \approx \sigma^4$, then $\dot{\mu}_3 = 6\mu\mu_3$ and $\mu_3(t) \equiv 0$. Equation (24) becomes

$$\left. \begin{aligned} \frac{d\mu}{dt} &= \mu^2 + \sigma^2 \\ \frac{d\sigma^2}{dt} &= 4\mu\sigma^2 \end{aligned} \right\} \quad (25)$$

These equations have been solved for $\mu(t)$ by J. F. Jordan (see previous article, Section B) for the case α uniform on a given interval. The solution for $\mu(t)$ is reasonably close to the exact solution, which means that the assumption $\mu_4 \approx \sigma^4$ is reasonable.

The results here were obtained by means of the equation of continuity; Jordan obtained the same results by a different method.

D. An n Thrustor Reliability Problem for Electric Propulsion, W. Kizner

1. Introduction

In this article formulas are presented for the calculation of the mean and variance of the position and velocity of an idealized spacecraft in one-dimensional flight while being continuously accelerated by an electrically powered engine having an array of n thrustors. Variations in the thrust levels of each thrustor and in the mass flow rates will be neglected here. These variations can be modeled by white noise and the resulting variances, which are independent of those computed here, can be calculated by standard Kalman filters.

Let the acceleration be denoted by a . Then, assuming that n thrustors are functioning,

$$a = \frac{\sum_{i=1}^n F_i}{m} \quad (1a)$$

$$m = m_0 - \left[\sum_{i=1}^n \int_0^t \beta_i d\tau \right] \quad (1b)$$

where

F_i = force from i th thrustor

m = mass of spacecraft

m_0 = initial mass of spacecraft

β_i = mass flow rate

We assume that all F_i are equal to F and all β_i are equal to β , which is assumed to be constant here.

We next assume that the failure of any thrustor is independent of the others. In addition we assume that the probability that a working thrustor fail is independent of the time, so that for a time interval Δt the probability of failure is equal to $\alpha \Delta t$. If we assume that the probability is one, it can easily be shown that a thrustor is operating at time zero; then at time t the probability that it is operating, $\theta(t)$, is

$$\theta(t) = \exp(-\alpha t) \quad (2)$$

If we have n thrustors which are operating at time $t = 0$ and all have the same law of failure given by Eq. (2), then the probability that exactly ℓ of them are operating at time t is given by the binomial law

$$B(n, \ell, t) = \binom{n}{\ell} [\theta(t)]^\ell [1 - \theta(t)]^{n-\ell} \quad (3)$$

The mean value of the number that is operating and its variance is known to be

$$E\ell = n\theta(t) \quad (4)$$

$$\text{var } \ell = n\theta(t) [1 - \theta(t)] \quad (5)$$

2. Autocovariance of the Number of Working Thrustors and Mean Values for Derived Quantities

We shall assume that the changes in $a(t)$ are small so that we can use the result from linear systems theory that the expected value of the output (the acceleration, speed, etc.) is obtained by passing the expected value of the input (the sum of the thrust from each thrustor and the sum of the mass flow rate) through the system (Eq. 1). Thus,

$$\bar{a}(t) = \frac{n\theta(t)F}{m_0 - E \left[\sum_{i=1}^n \int_0^t \beta_i h_i(\tau) d\tau \right]} \quad (6)$$

where $h_i(\tau)$ is a function that is unity when the i th thrustor is working and zero otherwise.

$$\begin{aligned} E \left[\sum_{i=1}^n \int_0^t \beta_i h_i(\tau) d\tau \right] &= n \int_0^t \theta(\tau) \beta d\tau \\ &= n \int_0^t \exp(-\alpha\tau) \beta d\tau \\ &= \frac{n\beta}{\alpha} [1 - \exp(-\alpha t)] \end{aligned} \quad (7)$$

Substituting Eq. (7) into Eq. (6),

$$\bar{a}(t) = \frac{n\theta(t)F}{m_0 - \frac{n\beta}{\alpha} [1 - \theta(t)]} \quad (8)$$

We next consider the variance of a . From Eq. (1)

$$\delta a(t) = a(t) - \bar{a}(t) = \frac{\sum_{i=1}^n F_i}{m} - \bar{a}(t) \quad (9)$$

Using a first-order approximation,

$$\delta a(t) = \frac{[\ell(t) - n\theta(t)]F}{\bar{m}(t)} - \frac{n\theta(t)F\delta m(t)}{[\bar{m}(t)]^2} \quad (10)$$

where $\ell(t)$ is the number of thrusters working at time t ,

$$\bar{m}(t) = m_0 - \frac{n\beta}{\alpha} [1 - \theta(t)] \quad (11)$$

$$\delta m(t) = -\beta \int_0^t [\ell(\tau) - n\theta(\tau)] d\tau \quad (12)$$

Define

$$\delta \ell(t) = \ell(t) - n\theta(t) = \ell(t) - \bar{\ell}(t) \quad (13)$$

We wish to find $E[\delta \ell(t_1) \delta \ell(t_2)]$.

Assume that $t_2 > t_1$, and denote $\ell(t_1)$ by ℓ_1 and $\ell(t_2)$ by ℓ_2 . Then the conditional probability of ℓ_2 given ℓ_1 is from Eq. (3),

$$p(\ell_2 | \ell_1) = \binom{\ell_1}{\ell_2} [\theta(t_2 - t_1)]^{\ell_2} [1 - \theta(t_2 - t_1)]^{t_1 - t_2} \quad (14)$$

where $0 \leq \ell_2 \leq \ell_1$.

Thus, the joint probability

$$p(\ell_1, \ell_2) = \binom{n}{\ell_1} [\theta(t_1)]^{\ell_1} [1 - \theta(t_1)]^{n - \ell_1} \binom{\ell_1}{\ell_2} [\theta(t_2 - t_1)]^{\ell_2} \times [1 - \theta(t_2 - t_1)]^{t_1 - t_2} \quad (15)$$

Let

$$V(t_1, t_2) = E[\delta \ell_1 \delta \ell_2] = E[\ell_1 \ell_2] - \bar{\ell}_1 \bar{\ell}_2 \quad (16)$$

We will use the well-known result (based on Fubini's theorem)

$$E[\ell_1 \ell_2] = E\{\ell_1 [E(\ell_2 | \ell_1)]\} \quad (17)$$

Thus,

$$\begin{aligned} E(\ell_1 \ell_2) &= E\{\ell_1 [\ell_1 \theta(t_2 - t_1)]\} \\ &= \theta(t_2 - t_1) E\ell_1^2 \\ &= \theta(t_2 - t_1) \{n\theta(t_1) [1 - \theta(t_1)] + n^2 [\theta(t_1)]^2\} \end{aligned} \quad (18)$$

Substituting Eq. (18) into Eq. (16) and recognizing that

$$\theta(t_2) = \theta(t_1) \theta(t_2 - t_1) \quad (19)$$

we have for the autocovariance function for the number of working thrusters

$$V(t_1, t_2) = \theta(t_2 - t_1) \{n\theta(t_1) [1 - \theta(t_1)]\}, \quad t_2 > t_1 \quad (20)$$

When $t_1 = t_2$, Eq. (20) reduces to Eq. (5).

3. Statistical Description of Linear Transformations of the Variations With the Number of Working Thrusters

The method that has been chosen here to calculate such quantities as the variance of the speed and distance as a function of time depends on a matrix approximation to the linearized operator.

For instance, consider Eq. (12) for $\delta m(t)$. We can write this approximately in matrix form as

$$\begin{bmatrix} \delta m(t_1) \\ \delta m(t_2) \\ \delta m(t_3) \\ \vdots \\ \delta m(t_i) \end{bmatrix} = -\beta \Delta t \begin{bmatrix} 0 & & & & \\ \frac{1}{2} & \frac{1}{2} & & & \\ & \frac{1}{2} & 1 & \frac{1}{2} & \\ & & \vdots & \ddots & \\ \frac{1}{2} & 1 & 1 & \cdots & \frac{1}{2} \end{bmatrix} \begin{bmatrix} \delta \ell_1 \\ \delta \ell_2 \\ \delta \ell_3 \\ \vdots \\ \delta \ell_i \end{bmatrix} \quad (21)$$

where Δt is the step size, or $t_2 - t_1, t_3 - t_2, \dots, t_{i+1} - t_i$. Thus, Eq. (21) is merely the trapezoidal rule applied to the quadrature represented by Eq. (12). As Δt approaches zero, Eq. (21) will tend toward the exact solution. For the other quantities of interest, we have

$$\delta a_i = \frac{\delta \ell(t_i) F}{\bar{m}(t_i)} + \frac{\beta n \theta(t_i) F}{[\bar{m}(t_i)]^2} \left[\sum_{j=0}^i \delta \ell(t_j) \Delta t \right] \quad (22)$$

where $\delta a_i = \delta a(t_i)$ and the prime on the sigma means that the first and last term in the summation is multiplied by $\frac{1}{2}$.

$$\delta v_0 = \Delta t \sum_{j=0}^i \delta a_j \quad (23)$$

$$\delta x_i = \Delta t \sum_{j=0}^i (t_i - t_j) \delta a_j \quad (24)$$

These equations then relate the change of δl to the quantities of interest. Let T represent the transformation matrix

$$y = Tl \quad (25)$$

where y is any variable of interest. Then the variance of y is given by

$$E[(y - \bar{y})(y^T - \bar{y}^T)] = T\Lambda T^T \quad (26)$$

where Λ is the autocovariance function of l .

4. Numerical Results

The following set of parameters were suggested by an asteroid belt rendezvous mission. Since there are so many idealizations in this one-dimensional problem, the numerical results given here can only be used in a very qualitative way.

Initial total spacecraft mass m_0 , kg	1014
Propellant mass used, kg	155
Total time of flight, days	400
Hyperbolic excess velocity, km/s	3.503
Effective jet velocity, km/s	50

If we assume a constant mass flow rate, then $\beta = 0.4485 \times 10^{-3}$ kg/s. The initial acceleration is 0.0002212 m/s.

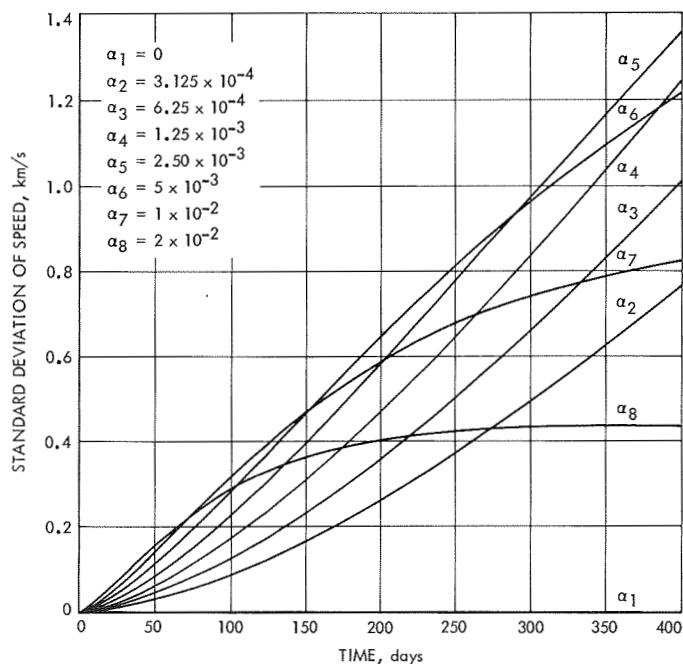


Fig. 3. Standard deviation of speed as a function of time

The main results are contained in Figs. 3 and 4, showing the standard deviation of speed and position, and in Figs. 5 and 6, showing the differences of average speed and position from the case where α is zero, or no

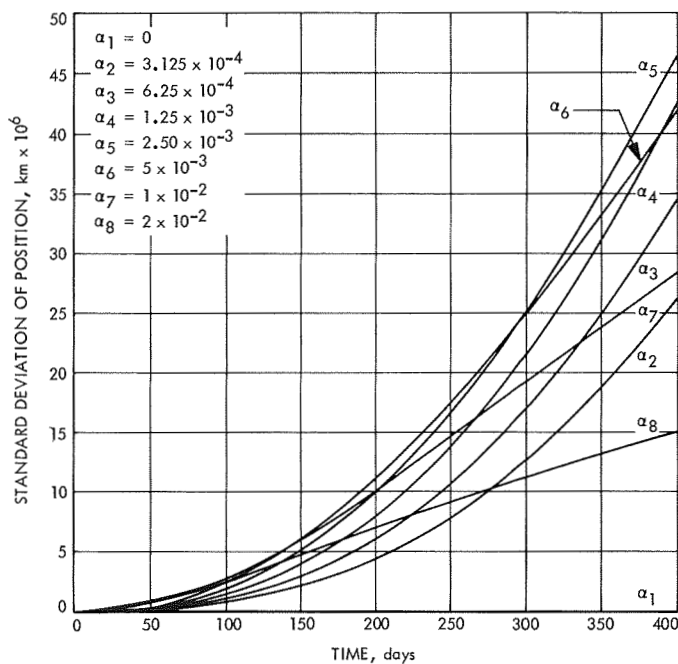


Fig. 4. Standard deviation of position as a function of time

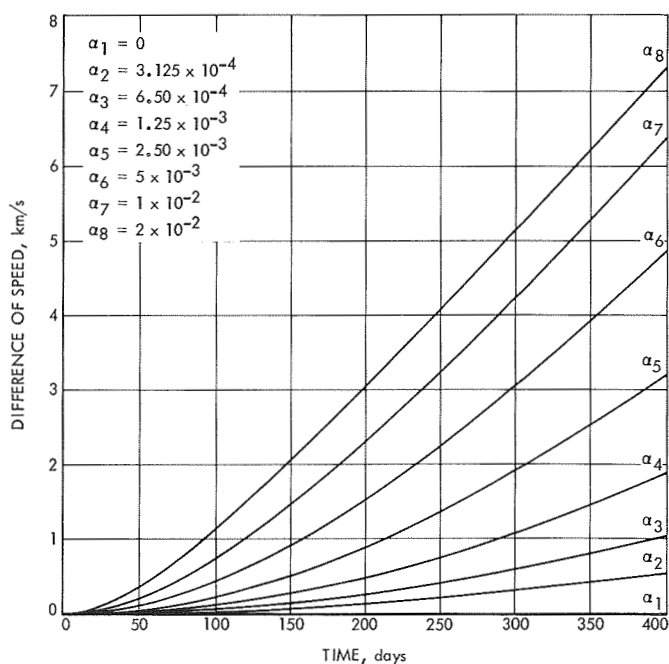


Fig. 5. Difference of speed as a function of time

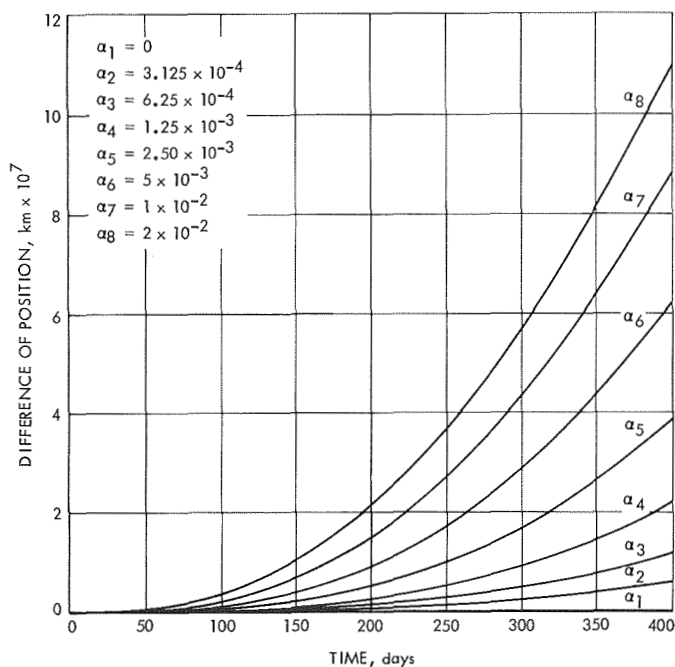


Fig. 6. Difference of position as a function of time

failures occur. The units are kilometers for distance and kilometers per second for speed. The unit for alpha, the failure rate constant, is failures per day, or the reciprocal of the half life of a thruster.

In Fig. 3, we see how increasing alpha from zero results in an increase in the standard deviation of the speed, until

some particular alpha value which depends on the time. Thereafter increasing alpha decreases the standard deviation. This may be qualitatively understood by considering that alpha equal to zero (no failures) produces zero standard deviation, while an infinite alpha would result in no motion and therefore in zero standard deviation. We also see that for a value of alpha of 0.0025, or 1/400 days, the standard deviation is approximately a maximum at the longer times. A similar situation occurs in the case of the standard deviation of position, as seen in Fig. 4.

Figures 5 and 6 show the differences in average speed and position from the case where no thruster failures occur. It is interesting to note that these differences are about the order of the corresponding standard deviations.

It is interesting to compare these standard deviations with a bias in the thrust level of the engine. If we imagine a bias in the acceleration with standard deviation σ_a ,

$$\sigma_v = \sigma_a t$$

and

$$\sigma_x = \frac{1}{2} \sigma_a t^2$$

A change in acceleration of 1% of the initial value (i.e., 0.2212×10^{-5} m/s) results in $\sigma_v = 0.076$ km/s and $\sigma_x = 1.31 \times 10^6$ km, which are considerably smaller values than those obtained from the analysis of thruster failure.